

Université Lumière Lyon 2

École Doctorale Humanité et Sciences Humaines

Faculté des Lettres, Sciences du Langage et Arts

Discipline : Sciences du Langage

Thèse pour obtenir le grade de Docteur en Sciences du Langage

Option : Phonétique/phonologie

Indices dynamiques et perception des voyelles :

Étude translinguistique en arabe dialectal et en français

Présentée et soutenue publiquement par

Jalal-eddin AL-TAMIMI

Le 14 Septembre 2007

Sous la direction de Monsieur Jean-Marie HOMBERT

Devant un jury composé de :

M. Thami BENKIRANE	Professeur à l'Université Sidi Mohammed Ben Abdellah-Fès, Maroc.	Examineur
M. René CARRÉ	Directeur de Recherche Émérite - CNRS et Université Lyon 2, France.	Examineur
M. Salem GHAZALI	Professeur à l'Université de 7 Novembre à Carthage, Tunisie.	Rapporteur
M. Jean-Marie HOMBERT	Directeur de Recherche – CNRS et Université Lyon 2, France.	Directeur de thèse
M. Björn LINDBLOM	Professeur Émérite à l'Université de Stockholm, Suède.	Rapporteur
M. Willy SERNICLAES	Chargé de Recherche 1^{er} classe et HDR, Université Paris 5, France.	Examineur

Laboratoire Dynamique du Langage – UMR 5596

Je dédie cette thèse à toute ma famille et à Loana.

REMERCIEMENTS

Je tiens tout d'abord à remercier Monsieur Jean-Marie HOMBERT d'avoir accepté de diriger ma thèse et de m'avoir aidé à choisir la voie scientifique que je suis actuellement.

Je voudrais également remercier tous les membres du jury pour avoir accepté ce rôle et pour l'honneur qu'ils me font de juger mon travail.

Je remercie tout particulièrement Monsieur René CARRÉ, sans l'aide de qui ce travail n'aurait pas vu le jour. Je vous remercie du fond du cœur pour les conseils méthodologiques, bibliographiques et de recherche. Je vous remercie également pour l'élaboration des différents programmes informatiques qui m'ont permis d'effectuer les expériences de perception et d'avoir à maintes reprises accepté de corriger à la dernière minute un bug informatique, et de m'avoir lu et relu jusqu'à tard le soir. De m'avoir conseillé pour l'organisation de ma thèse.

Je tiens à remercier également Monsieur Salem GHAZALI pour les conseils qu'il m'a donnés pour ma recherche.

Je tiens à exprimer ma grande gratitude envers Monsieur Thami BENKIRANE qui m'a aidé dans l'élaboration du second corpus en arabe marocain, de m'avoir accueilli chez lui à Fès et de m'avoir ouvert les portes de l'université pour trouver les locuteurs qui ont participé à cette expérience.

Je tiens à remercier également Monsieur François PELLEGRINO le directeur du Laboratoire Dynamique du Langage, qui a accepté de m'aider à plusieurs reprises pour l'élaboration des interfaces des programmes de passation des expériences en production et en perception, les différents outils qui me permettaient d'effectuer quelques calculs mathématiques poussés. Surtout d'avoir accepté de me relire plusieurs fois, de me conseiller, de me critiquer, de me refaire faire une bonne partie de ma thèse, ;-), mais tout simplement d'être présent quand j'avais de réels problèmes et de m'avoir trouvé des solutions.

Je ne peux pas passer à côté du grand DON RIGATONI (alias Egidio MARSICO) qui m'a toujours inspiré par son humour, sa capacité à être là, de m'avoir énormément écouté dans tous les moments difficiles ou tout simplement pour rigoler, de m'avoir lu, relu, et encore relu plusieurs fois, de m'avoir toujours

conseillé de faire ceci ou cela autour d'une recette de *Pastasciutta*, ou d'un très bon café très serré à l'italienne comme on dit. Je te remercie du fond du cœur d'être là et d'avoir tout fait pour faire naître ce travail. Sans toi, je ne pouvais rien faire. Surtout refaire presque toute ma thèse !!!

Je tiens également à remercier Manu (alias Emmanuel FERRAGNE), notre cher représentant des doctorants et responsable des cotisation de café, et surtout mon colocataire du bureau du fond du couloir (oui tout au fond). Je te remercie d'avoir été là devant moi pendant deux ans maintenant et d'avoir été toujours à l'écoute de mes questions qui démarraient tout le temps par, « *hé, manu, je peux te poser une question, blablabla...* » ou sinon c'était mon petit regard qui te faisait savoir que j'allais poser une question. Merci de m'avoir initié à Praat et à MatLab et de m'avoir beaucoup aidé pour démarrer l'écriture des scripts. Merci pour toutes les discussions qu'on a eues concernant les pentes et les transitions formantiques, les systèmes vocaliques, la lecture de notre vieux spectro. Merci de m'avoir supporté quand je sifflotais tout le temps sans faire attention. Merci beaucoup pour tout ce que tu m'as apporté.

Je voudrais également remercier tous les membres du Laboratoire Dynamique du Langage qui m'ont accepté parmi eux dès ma maîtrise. Tout particulièrement, je voudrais remercier notre grand informaticien Christian FRESSARD sans l'aide de qui tout doctorant et chercheur ne peut faire marcher son ordinateur. D'être là très tôt le matin prêt à nous aider et à écouter à fond sa musique et de m'avoir aidé quand j'avais plein de problèmes informatiques entre plantage de l'ordinateur et jusqu'à récemment le grillage de mon disque dur. Je voudrais remercier Linda BRENDELIN pour toute l'aide qu'elle nous apporte au niveau de la gestion des missions, de l'organisation administrative du laboratoire, d'avoir accepté toutes les discussions qu'on a eues sur tout type de sujet : problèmes administratifs, organisation de voyage, ou tout simplement recherche d'appartement, etc... Je voudrais également remercier Vincent MONATTE qui m'a beaucoup aidé avec les cartes, les discussions qu'on a eues. Je ne dois pas oublier non plus de remercier Christof dosh (alias Christophe DOS SANTOS) qui a démarré sa thèse en même temps que moi et a eu à plusieurs reprises des discussions fructueuses avec moi autour d'une tasse de thè(se) et de café et surtout d'un grand cru de chocolat. Tu m'as beaucoup aidé, merci. Je remercie donc toutes les personnes présentes au laboratoire et plus

particulièrement : Nathalie Bedoin (qui m'a beaucoup aidé, conseillé et surtout lu) ; Gérard Philippson (pour ses conseils) ; Christophe Coupé ; les deux Claire (Grataloup et Delle Luche), Caroline Jacquier, Rim Hamdi, Laëtitia Savot et Marion Girard (mes anciennes colocataires de bureau 206C), Florence Chenu, Arnaud Sicard, Liliane, Taoufiq (qui ne fait pas partie du labo, je sais !!) et tous ceux que j'ai oubliés. Je n'oublie pas non plus les grandes pauses café-jeu juste après le repas de midi, où une initiation à l'art de Banzaï, de Kamelot, de Non Merci et des Vaches a été proposée presque tous les jours. Je n'oublie pas non plus les bons moments qu'on passe souvent ensemble surtout le jeudi entre 18h30 et 20h pour nous initier à l'art de « *jouer comme des malades* » où on peut facilement se tordre une cheville, ou se faire casser ses lunettes, ou tout simplement s'amuser à marquer des paniers à 10 points (hé scude !!!).

Je voudrais également remercier Mohamed Embraki qui m'a beaucoup aidé et conseillé ces derniers temps pour la partie arabe et la partie sur la coarticulation. Je remercie également Christelle Dodane, qui m'a apporté tout son soutien.

Je voudrais surtout remercier les 60 locuteurs qui ont participé aux différentes expériences, sans l'aide de qui ce travail n'aurait pas vu le jour.

Je remercie toute ma famille en Jordanie : Jamil, Mozayyan, Jamal, Meena, Maysa, Souma, Ousama et tous mes amis jordaniens qui sont restés à mes côtés dès mon arrivée en France.

Pour terminer ces longs remerciements, je te remercie Loana, d'être présente à mes côtés et de m'avoir aidé et soutenu jusqu'au bout dans les bons et les mauvais moments. Sans toi je n'aurais pas pu continuer...

TABLE DES MATIÈRES

REMERCIEMENTS	I
TABLE DES MATIÈRES	V
CONVENTIONS.....	XIII
NORMES DES RÉFÉRENCES BIBLIOGRAPHIQUES.....	XIII
CONVENTIONS DE TRANSCRIPTION DE L'ARABE.....	XIV
CONVENTIONS DE TRANSCRIPTION PHONÉTIQUE	XV
INTRODUCTION GÉNÉRALE	1
CHAPITRE 1. LES SYSTÈMES VOCALIQUES DE L'ARABE	11
1.1. INTRODUCTION	11
1.2. LA VOYELLE EN ARABE.....	15
1.2.1. <i>Vision des grammairiens arabes</i>	16
1.2.2. <i>Les études actuelles</i>	19
1.3. CARACTÉRISTIQUES GÉNÉRALES DE LA LANGUE ARABE	24
1.3.1. <i>Histoire de la langue Arabe</i>	25
1.3.2. <i>Structure de la langue arabe</i>	29
1.3.3. <i>Typologie des dialectes arabes</i>	31
1.3.3.1. Division Géographique.....	31
1.3.3.2. Division Sociolinguistique	38
1.4. DESCRIPTION DES DEUX DIALECTES ARABES ÉTUDIÉS	43
1.4.1. <i>Arabe Jordanien</i>	43
1.4.1.1. Généralités.....	43
1.4.1.2. Le parler d'Irbid.....	48
1.4.2. <i>Arabe Marocain</i>	50
1.4.2.1. Généralités.....	50
1.4.2.2. Le parler de Casablanca.....	53
1.4.2.3. Le parler de Fès	54
1.5. CONCLUSIONS.....	55
CHAPITRE 2. LES MODÈLES DE PERCEPTION DES VOYELLES ET DE LA PAROLE.....	59
2.1. INTRODUCTION	59
2.2. LES APPROCHES EN PERCEPTION DES VOYELLES	60
2.2.1. <i>L'approche « Simple Target »</i>	60
2.2.2. <i>L'approche « Elaborated Target »</i>	63
2.2.3. <i>L'approche « Dynamic Specification »</i>	68

2.3.	THÉORIES DE PERCEPTION DE LA PAROLE	77
2.3.1.	<i>Théories Passives</i>	78
2.3.2.	<i>Théories Actives</i>	80
2.3.2.1.	La Théorie Motrice de perception de la parole	82
2.3.2.2.	Le modèle de l'Analyse par Synthèse	86
2.3.2.3.	La Théorie Quantique.....	88
2.3.2.4.	La Théorie H&H	91
2.3.2.5.	Le rôle des « hidden gestures » en Phonologie Articulatoire.....	95
2.3.2.6.	Le Native Language Magnet model	97
2.3.2.7.	Le prototype & « The Hyper-Space effect ».....	103
2.4.	CONCLUSIONS.....	106
CHAPITRE 3. PROBLÉMATIQUE		109
CHAPITRE 4. RÉSULTATS EN PRODUCTION DES VOYELLES		117
4.1.	INTRODUCTION	117
4.2.	MÉTHODOLOGIE DU TRAVAIL	118
4.2.1.	<i>Locuteurs</i>	118
4.2.2.	<i>Corpus</i>	119
4.2.3.	<i>Protocole expérimental en production</i>	120
4.2.4.	<i>Segmentation du corpus</i>	122
4.3.	RÉSULTATS INTER-LANGUES.....	125
4.3.1.	<i>Analyses statiques des formants des voyelles</i>	126
4.3.1.1.	Analyse et traitement des données.....	126
4.3.1.2.	Résultats de dispersion des espaces vocaliques.....	131
4.3.1.3.	Résultats des ellipses de dispersion de chaque catégorie vocalique	145
4.3.2.	<i>Analyses dynamiques des formants des voyelles</i>	149
4.3.2.1.	Analyse et traitement des données.....	149
4.3.2.1.1.	La transition formantique et le « locus virtuel ».....	150
4.3.2.1.2.	Les équations de locus.....	154
4.3.2.1.3.	Les pentes formantiques.....	159
4.3.2.2.	Résultats des pentes formantiques.....	165
4.3.2.2.1.	Sans normalisation du temps.....	167
4.3.2.2.2.	Avec normalisation du temps.....	175
4.3.3.	<i>Discussion des résultats inter-langues en statique vs en dynamique</i>	182
4.3.4.	<i>Résultats de l'analyse discriminante</i>	190
4.3.4.1.	En statique.....	190
4.3.4.2.	En dynamique.....	195
4.3.4.2.1.	Régression linéaire.....	195
4.3.4.2.2.	Régression polynomiale de 2 ^{ème} ordre.....	200
4.3.4.2.3.	Régression polynomiale de 3 ^{ème} ordre.....	204
4.4.	SPÉCIFICITÉS DE CHAQUE LANGUE.....	208
4.4.1.	<i>En Arabe Marocain</i>	208
4.4.1.1.	Résultats en statique	208

4.4.1.2.	Résultats en dynamique.....	218
4.4.1.3.	Résultats de l'analyse discriminante	223
4.4.1.3.1.	En statique.....	223
4.4.1.3.2.	En dynamique	227
4.4.1.3.2.1.	Régression linéaire.....	227
4.4.1.3.2.2.	Régressions polynomiales de 2 ^{ème} et 3 ^{ème} ordre	229
4.4.1.4.	Discussion des résultats en statique vs en dynamique	233
4.4.2.	<i>En Arabe Jordanien</i>	235
4.4.2.1.	Résultats en statique	236
4.4.2.2.	Résultats en dynamique.....	246
4.4.2.3.	Résultats de l'analyse discriminante	253
4.4.2.3.1.	En statique.....	253
4.4.2.3.2.	En dynamique	256
4.4.2.3.2.1.	Régression linéaire.....	256
4.4.2.3.2.2.	Régressions polynomiales de 2 ^{ème} et 3 ^{ème} ordre	259
4.4.2.4.	Discussion des résultats en statique vs en dynamique	263
4.4.3.	<i>En Français</i>	265
4.4.3.1.	Résultats en statique	267
4.4.3.2.	Résultats en dynamique.....	276
4.4.3.3.	Résultats de l'analyse discriminante	281
4.4.3.3.1.	En statique.....	281
4.4.3.3.2.	En dynamique	285
4.4.3.3.2.1.	Régression linéaire.....	285
4.4.3.3.2.2.	Régressions polynomiales de 2 ^{ème} et 3 ^{ème} ordre	287
4.4.3.4.	Discussion des résultats en statique vs en dynamique	291
4.5.	RÉCAPITULATIF DES RÉSULTATS EN STATIQUE VS EN DYNAMIQUE.....	293
4.6.	CONCLUSIONS.....	298
CHAPITRE 5. RÉSULTATS EN PERCEPTION DES VOYELLES		301
5.1.	INTRODUCTION	301
5.2.	MÉTHODOLOGIE DU TRAVAIL.....	303
5.2.1.	<i>Locuteurs</i>	303
5.2.2.	<i>Corpus</i>	303
5.2.3.	<i>Protocoles expérimentaux en perception</i>	305
5.2.3.1.	Recherche de prototypes vocaliques.....	306
5.2.3.2.	Catégorisation de l'espace vocalique.....	310
5.3.	RECHERCHE DE PROTOTYPES VOCALIQUES	314
5.3.1.	<i>Analyse et traitement des données</i>	314
5.3.2.	<i>En Arabe Marocain</i>	315
5.3.2.1.	Résultats en statique (V Target)	315
5.3.2.2.	Résultats en dynamique (CV Target).....	320
5.3.2.3.	Comparaison des résultats en statique vs en dynamique.....	324
5.3.2.4.	Résultats de l'analyse discriminante	331
5.3.3.	<i>En Arabe Jordanien</i>	335

5.3.3.1.	Résultats en statique (V Target)	336
5.3.3.2.	Résultats en dynamique (CV Target).....	339
5.3.3.3.	Comparaison des résultats en statique vs en dynamique.....	344
5.3.3.4.	Résultats de l'analyse discriminante	350
5.3.4.	<i>En Français</i>	355
5.3.4.1.	Résultats en statique (V Target)	355
5.3.4.2.	Résultats en dynamique (CV Target).....	359
5.3.4.3.	Comparaison des résultats en statique vs en dynamique.....	363
5.3.4.4.	Résultats de l'analyse discriminante	367
5.3.5.	<i>Résultats Inter-Langues</i>	372
5.3.5.1.	Résultats en statique (V Target)	372
5.3.5.2.	Résultats en dynamique (CV Target).....	379
5.3.5.3.	Comparaison des résultats en statique vs en dynamique.....	384
5.3.5.4.	Résultats de l'analyse discriminante	391
5.4.	CATÉGORISATION DE L'ESPACE VOCALIQUE.....	395
5.4.1.	<i>Analyses et traitement des données</i>	395
5.4.2.	<i>En Arabe Marocain</i>	396
5.4.2.1.	Résultats en statiques (V Target).....	396
5.4.2.2.	Résultats en dynamique (CV Target).....	401
5.4.3.	<i>En Arabe Jordanien</i>	406
5.4.3.1.	Résultats en statique (V Target)	406
5.4.3.2.	Résultats en dynamique (CV Target).....	410
5.4.4.	<i>En Français</i>	416
5.4.4.1.	Résultats en statique (V Target)	416
5.4.4.2.	Résultats en dynamique (CV Target).....	418
5.5.	CONCLUSIONS.....	420
CONCLUSION GÉNÉRALE ET PERSPECTIVES		425
BIBLIOGRAPHIE		441
RÉFÉRENCES CONSULTÉES		441
BIBLIOGRAPHIES CITÉES		457
ANNEXES	467	
ANNEXE 1	ESPACES VOCALIQUES ET AIRES DE DISPERSION DES VOYELLES EN ARABE MAROCAIN, ARABE JORDANIEN ET EN FRANÇAIS	469
ANNEXE 2	PENTES FORMANTIQUES SANS NORMALISATION DU TEMPS EN SITUATION SYLLABE EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS	474
ANNEXE 3	PENTES FORMANTIQUES SANS NORMALISATION DU TEMPS EN SITUATION ISOLATION EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS	477
ANNEXE 4	PENTES FORMANTIQUES AVEC NORMALISATION DU TEMPS EN SITUATION SYLLABE EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS	480

ANNEXE 5	PENTES FORMANTIQUES AVEC NORMALISATION DU TEMPS EN SITUATION ISOLATION EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS	483
ANNEXE 6	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN STATIQUE AVEC LA DURÉE VOCALIQUE (AD) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.....	486
ANNEXE 7	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN STATIQUE SANS LA DURÉE VOCALIQUE (SD) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.....	488
ANNEXE 8	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION LINÉAIRE) SANS NORMALISATION DU TEMPS (SN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.....	490
ANNEXE 9	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC NORMALISATION DU TEMPS (AN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.....	492
ANNEXE 10	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2 ^{ÈME} ORDRE) SANS NORMALISATION DU TEMPS (SN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.....	494
ANNEXE 11	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2 ^{ÈME} ORDRE) AVEC NORMALISATION DU TEMPS (AN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.	496
ANNEXE 12	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3 ^{ÈME} ORDRE) SANS NORMALISATION DU TEMPS (SN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.....	498
ANNEXE 13	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3 ^{ÈME} ORDRE) AVEC NORMALISATION DU TEMPS (AN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.	500
ANNEXE 14	PENTES FORMANTIQUES DE F ₁ , F ₂ & F ₃ PAR VOYELLE EN SITUATIONS SYLLABE & ISOLATION EN ARABE MAROCAIN	502
ANNEXE 15	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN STATIQUE AVEC LA DURÉE VOCALIQUE (AD) EN SITUATIONS SYLLABE ET ISOLATION.	508
ANNEXE 16	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN STATIQUE SANS LA DURÉE VOCALIQUE (SD) EN SITUATIONS SYLLABE ET ISOLATION.	510

ANNEXE 17	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.....	512
ANNEXE 18	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2 ^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.....	514
ANNEXE 19	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3 ^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.....	516
ANNEXE 20	PENTES FORMANTIQUES (F ₁ , F ₂ , F ₃) PAR VOYELLE EN SYLLABE & ISOLATION EN ARABE JORDANIEN	518
ANNEXE 21	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN STATIQUE AVEC LA DURÉE VOCALIQUE (AD) EN SITUATIONS SYLLABE ET ISOLATION.	524
ANNEXE 22	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN STATIQUE SANS LA DURÉE VOCALIQUE (SD) EN SITUATIONS SYLLABE ET ISOLATION.	526
ANNEXE 23	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.....	528
ANNEXE 24	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2 ^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.....	530
ANNEXE 25	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3 ^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.....	532
ANNEXE 26	VOYELLES INDIVIDUELLES EN FRANÇAIS SELON LES RÉGIONS : CENTRE (EN BLEU) ET SUD (EN ROUGE).....	534
ANNEXE 27	PENTES FORMANTIQUES DE F ₁ , F ₂ & F ₃ PAR VOYELLE EN SITUATIONS SYLLABE & ISOLATION EN FRANÇAIS.	537
ANNEXE 28	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN STATIQUE AVEC LA DURÉE VOCALIQUE (AD) EN SITUATIONS SYLLABE ET ISOLATION.	543
ANNEXE 29	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN STATIQUE SANS LA DURÉE VOCALIQUE (SD) EN SITUATIONS SYLLABE ET ISOLATION.	545

ANNEXE 30	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.....	547
ANNEXE 31	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2 ^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.	549
ANNEXE 32	MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3 ^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.	551
ANNEXE 33	ELLIPSES DE DISPERSION DES PROTOTYPES VOCALIQUES EN ARABE MAROCAIN EN FONCTION DU LIEU D'ARTICULATION.....	553
ANNEXE 34	ELLIPSES DE DISPERSION DES PROTOTYPES VOCALIQUES EN ARABE JORDANIEN EN FONCTION DU LIEU D'ARTICULATION.....	554
ANNEXE 35	ELLIPSES DE DISPERSION DES PROTOTYPES VOCALIQUES EN FRANÇAIS EN FONCTION DU LIEU D'ARTICULATION.....	555

CONVENTIONS

NORMES DES RÉFÉRENCES BIBLIOGRAPHIQUES

La plupart des références bibliographiques citées dans la thèse ont été consultées directement. Il n'a pas été possible de nous procurer certaines de ces références. Ainsi, nous avons décidé de marquer ces dernières par des astérisques après la date (e.g., Liberman *et al.*, 1967_b*) dans le corps de la thèse. Nous avons divisé la partie concernant les références bibliographiques en deux : la bibliographie consultée et la bibliographie citée dans laquelle nous avons systématiquement marqué la provenance.

CONVENTIONS DE TRANSCRIPTION DE L'ARABE

Nous avons utilisé le système de translittération de l'arabe proposé par *The Institute of Islamic Studies* de l'Université McGill au Canada. Le tableau suivant rend compte des correspondances entre les caractères arabes, les symboles phonétiques correspondants et les symboles utilisés pour la translittération.

Caractère arabe	Transcription en API	Translittération	Caractère arabe	Transcription en API	Translittération	Caractère arabe	Transcription en API	Translittération
ا	ʔ	ʔ	ع	ʃ	š	ن	n	n
ب	b	b	ط	sʔ	ʂ	هـ	h	h
ت	t	t	ظ	dʔ	ɖ	و	w	w
ث	θ	ṯ	ط	tʔ	ṯ	ي	j	y
ج	ʒ	gɟ	ظ	ðʔ	ʒ	آ	a:	ā
ح	ħ	ħ	ع	ʕ	ˤ	إ	i:	ī/ē
خ	χ	ħ	غ	ɣ	g̃	و	u:	ū/ō
د	d	d	ف	f	f	أ	a	a
ذ	ð	d̪	ق	q	q	إ	i	i
ر	r	r	ك	k	k	أ	u	u
ز	z	z	ل	l	l	ع	ʃʃ	šš
س	s	s	م	m	m			

CONVENTIONS DE TRANSCRIPTION PHONÉTIQUE

La transcription phonétique utilisée s'appuie sur les normes utilisées par l'Alphabet Phonétique International présenté dans le tableau qui suit.

Source : <http://www.arts.gla.ac.uk/IPA/fullchart.html>

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2005)

CONSONANTS (PULMONIC) © 2005 IPA

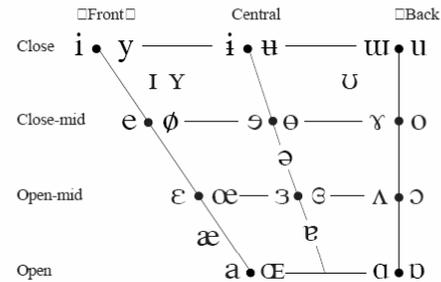
	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			ʀ					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
◌ ɸ Bilabial	ɓ Bilabial	ʼ Examples:
◌ ɗ Dental	ɗ Dental/alveolar	ɓʼ Bilabial
◌ ɗ̥ (Post)alveolar	ɟ Palatal	ɗʼ Dental/alveolar
◌ ɗ̥ Palatoalveolar	ɠ Velar	ɟʼ Velar
◌ ɗ̥ Alveolar lateral	ɠ Uvular	ɟʼ Alveolar fricative

VOWELS



Where symbols appear in pairs, the one to the right represents a rounded vowel.

OTHER SYMBOLS

ɱ Voiceless labial-velar fricative	ɕ ʑ Alveolo-palatal fricatives
ɰ Voiced labial-velar approximant	ɭ Voiced alveolar lateral flap
ɰ Voiced labial-palatal approximant	ɥ Simultaneous ʃ and X
ħ Voiceless epiglottal fricative	
ʕ Voiced epiglottal fricative	Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary.
ʔ Epiglottal plosive	

kp̥ ts̥

SUPRASEGMENTALS

- ˈ Primary stress
- ˌ Secondary stress
- ː Long *founəˈtʃən*
- ˑ Half-long *eˑ*
- ˘ Extra-short *e˘*
- ◌ Minor (foot) group
- ◌ Major (intonation) group
- ◌ Syllable break *ti.ækt*
- ◌ Linking (absence of a break)

DIACRITICS Diacritics may be placed above a symbol with a descender, e.g. ɨ̥̄

◌ Voiceless <i>n̥ d̥</i>	◌ Breathy voiced <i>b̤ a̤</i>	◌ Dental <i>t̪ d̪</i>
◌ Voiced <i>ṅ ṭ</i>	◌ Creaky voiced <i>b̰ a̰</i>	◌ Apical <i>t̺ d̺</i>
◌ Aspirated <i>tʰ dʰ</i>	◌ Linguolabial <i>t̟ d̟</i>	◌ Laminal <i>t̟ d̟</i>
◌ More rounded <i>ɔ̹</i>	◌ Labialized <i>tʷ dʷ</i>	◌ Nasalized <i>ẽ̃</i>
◌ Less rounded <i>ɔ̜</i>	◌ Palatalized <i>tʲ dʲ</i>	◌ Nasal release <i>d̪ⁿ</i>
◌ Advanced <i>u̟</i>	◌ Velarized <i>t̟̠ d̟̠</i>	◌ Lateral release <i>d̪ˡ</i>
◌ Retracted <i>e̠</i>	◌ Pharyngealized <i>t̟̠̠ d̟̠̠</i>	◌ No audible release <i>d̪̠</i>
◌ Centralized <i>ẽ̜</i>	◌ Velarized or pharyngealized <i>ɫ</i>	
◌ Mid-centralized <i>ẽ̞</i>	◌ Raised <i>e̝</i> (<i>ɹ̝</i> = voiced alveolar fricative)	
◌ Syllabic <i>n̩</i>	◌ Lowered <i>e̞</i> (<i>β̞</i> = voiced bilabial approximant)	
◌ Non-syllabic <i>e̯</i>	◌ Advanced Tongue Root <i>e̟̟̟</i>	
◌ Rhoticity <i>ə̣ ɑ̣</i>	◌ Retracted Tongue Root <i>e̠̠̠</i>	

TONES AND WORD ACCENTS

- | LEVEL | CONTOUR |
|---------------------|---------------------------|
| é̋ or ɛ̋ Extra high | é̋ or ɛ̋ / Rising |
| é̌ or ɛ̌ High | é̌ or ɛ̌ \ Falling |
| é̍ or ɛ̍ Mid | é̍ or ɛ̍ / High rising |
| é̎ or ɛ̎ Low | é̎ or ɛ̎ / Low rising |
| é̏ or ɛ̏ Extra low | é̏ or ɛ̏ / Rising-falling |
| ↓ Downstep | ↘ Global rise |
| ↑ Upstep | ↙ Global fall |

INTRODUCTION GÉNÉRALE

« *Experimental findings show that perception is more sensitive to changing stimulus patterns than to purely static ones. In speech articulatory movement and spectral changes are pervasive whereas steady-state intervals are rare. In view of those observations, it may seem curious that the traditional frameworks for describing speech should tend to emphasize static attributes. For instance, the IPA framework specifies consonants and vowels using terms the majority of which refer to speech production steady-states, e.g., labial, high, voiced, stop, nasal, etc...*

If perception likes change, why is phonetic specification built mainly around steady-state attributes? ».

(Lindblom et al., 2006 : pp. 7-8).

Quelle est la nature des indices identifiables en production et utilisés en perception ? Ces indices correspondent-ils à des positions articulatoires et acoustiques **statiques** et stables dans le conduit vocal, comme le définissent les partisans de la théorie « *Simple Target* » (Joos, 1948 ; Ladefoged, 1967 ; Daniloff & Hammarberg, 1973 ; etc...), ou sont-ils plutôt de nature **dynamique**, comme le définissent les partisans de la théorie « *Dynamic Specification* » (Fairbanks & Grubb, 1961* ; Fujimura & Ochiai, 1963 ; Lehiste & Meltzer, 1973 ; Strange *et al.*, 1976, 1983 ; Strange, 1989, 1999_a ; Carré, 2004, 2006, 2007 (à paraître) ; Carré *et al.*, 2004 ; Lindblom *et al.*, 2006 ; etc...). Enfin, les indices utilisés en perception peuvent être multiples et donc à la fois **statiques** et **dynamiques** (Carré, 2004, 2006, 2007 (à paraître) ; Carré *et al.*, 2004 ; Strange 1989, 1999_a ; etc...) ? Plus spécifiquement, quelle unité perceptive minimale permet d'identifier les voyelles : est-ce une « cible vocalique » (ou les *V Target*) résultant d'une configuration **statique** ou un ensemble de traits **dynamiques**, qui prennent en compte les changements spectraux propres à la voyelle et/ou les effets coarticulaires des consonnes et/ou voyelles adjacentes (ou les *CV Target*) ?

La recherche de réponses à ces questions est le sujet principal de notre thèse où nous cherchons à évaluer le rôle d'indices **dynamiques** en perception, indices identifiés à partir des résultats de production et ce pour des voyelles de l'arabe dialectal et du français. Cette thèse s'inscrit dans la continuité des recherches que nous avons menées pour notre mémoire de DEA et qui avaient pour but la description des systèmes vocaliques de l'arabe dialectal jordanien et marocain ainsi que l'étude des rapports existant entre les voyelles produites et

perçues par les locuteurs natifs de ces deux dialectes en comparaison avec le français (voir Al-Tamimi, 2002 ; Al-Tamimi & Barkat-Defradas, 2002 ; Barkat-Defradas *et al.*, 2003¹). La tâche des locuteurs était, d'une part de produire les voyelles de leur langue au sein de mots, de syllabes et en isolation et, d'autre part de rechercher les prototypes vocaliques perceptifs correspondant à ces voyelles (via une méthode d'ajustement de formants). Lors de cette étude, nous nous sommes aperçu que les locuteurs arabophones avaient de grandes difficultés à produire et à percevoir des voyelles isolées, comparés aux locuteurs français. La plupart du temps, les locuteurs jordaniens produisaient les voyelles isolées de leur système comme des voyelles se rapprochant soit d'une voyelle centrale [ə] soit d'une voyelle ouverte [a]. En perception, les auditeurs arabophones exprimaient aussi leurs grandes difficultés à identifier les voyelles, car les sons qu'ils entendaient correspondaient à des « **bips** » ou à des « **bruits d'ordinateur** ».

Nous avons donc émis plusieurs hypothèses pouvant expliquer le comportement des locuteurs arabophones, parmi lesquelles : i) la qualité des sons de synthèse, ii) l'utilisation de stimuli **statiques** (i.e., suivant l'approche « *Simple Target* », voir *infra*), ou iii) l'organisation de la morphologie du lexique différents dans les langues.

Nous avons utilisé de la parole synthétisée pour tester nos hypothèses. Cet outil est d'une grande importance, car il permet d'obtenir toutes les combinaisons possibles des fréquences formantiques qui caractérisent les voyelles à l'intérieur de l'espace vocalique ; combinaisons impossibles à obtenir à partir de la production humaine. Or, différentes études ont montré que les auditeurs naïfs, rencontrent des difficultés liées à la qualité de ces sons synthétiques, en comparaison avec l'utilisation de la parole naturelle (Johnson *et al.*, 1999 ;

¹ Ce travail faisait partie d'un projet plus large « *Variabilité phonétique en production et perception de la parole: rôle et limites des stratégies individuelles* », financé par le Ministère de la Recherche et de l'Enseignement Supérieur. Ce projet avait pour but l'étude des liens entre la production et la perception de la parole au sein de trois langues ayant des systèmes vocaliques différents : le français (avec onze voyelles orales et représenté par les parlers du nord, du centre et du sud), l'italien (avec sept voyelles et représenté par le parler de Florence) et l'arabe (avec le dialecte jordanien d'Irbid (huit voyelles) et le dialecte marocain de Casablanca (avec cinq voyelles)). Ce projet visait avant tout à caractériser la variabilité individuelle afin d'étudier les processus de production et de perception des voyelles et à étudier les possibles stratégies individuelles qui peuvent la soutenir (Carré & Hombert, 2002).

Johnson, 2000). En nous inspirant des travaux de (Lieberman *et al.*, 1967_a, 1967_b*; Lieberman, 1970), nous considérons que les stimuli synthétiques qui ne présentent que des valeurs formantiques stables ne sont que des **artefacts de laboratoire**, ceci étant appuyé par des résultats perceptifs indiquant que les auditeurs éprouvent de grandes difficultés lors de l'identification de ces voyelles. Ceci a permis de développer les nouveaux courants en perception des voyelles et plus particulièrement les théories dites **dynamiques**.

Pourtant, les résultats obtenus pour la population française n'ont montré aucune difficulté liée à la tâche (i.e., ni pour les sons synthétiques, ni pour l'utilisation des stimuli **statiques**). Nous nous sommes alors retourné vers une explication plus structurelle. Nous savons que l'arabe d'un côté et le français et l'italien de l'autre, sont des langues différentes sur le plan morphologique : la langue arabe est une langue non-concaténative, où la racine consonantique trilitère joue un rôle déterminant dans la morphologie du lexique (McCarthy, 1979 ; Pierrehumbert, 1992, etc...) ; à l'opposé, le français et l'italien se caractérisent par une structure à dominance concaténative basée sur une morphologie flexionnelle. Nous pensons donc que la structure morphologique différente entre ces deux couples de langues peut être à l'origine d'une utilisation différente des voyelles et par conséquent d'un statut et d'un rôle différents de ces voyelles.

En effet, en ce qui concerne la langue arabe, les grammairiens arabes expliquent que, pour cause de structure à racine consonantique, les voyelles se superposent aux consonnes pour en faciliter la production ; la voyelle n'a ainsi qu'un rôle purement grammatical : la consonne transmet le sens lexical d'un mot tandis que la voyelle sert à définir la catégorie grammaticale d'un mot (Versteegh, 2001 ; Ryding, 2005). Un autre effet de l'influence de la morphologie en arabe est que la voyelle n'est jamais présente en isolation ou en début de syllabe et pour qu'elle soit réalisée, elle doit toujours être associée à une consonne ou à défaut, à un coup de glotte. Tout ceci a motivé certains auteurs à dire que la voyelle en arabe n'a peut-être pas de « représentation phonologique » simple chez les locuteurs arabophones (voir Cantineau, 1960, etc...).

Ceci nous conduit à penser que, puisque la voyelle en arabe ne possède peut-être pas de « représentation phonologique » simple, son identification par les auditeurs arabophones pourrait poser problème. Nous pouvons penser en effet qu'une simple caractérisation **statique** de la voyelle en arabe n'est pas pertinente, mais que c'est la prise en compte des influences consonantiques, et donc de l'entourage auquel elle est associée, qui facilitera sa perception par les auditeurs arabophones.

En fait, le débat autour de la représentation des voyelles en tant qu'éléments **statiques** vs **dynamiques** n'est pas nouveau, puisqu'il a alimenté les recherches en production et en perception de la parole depuis les années 60. En s'appuyant sur les « **cibles vocaliques** » produites en isolation ou en contexte, différents auteurs ont élaboré plusieurs modèles de perception des voyelles. Il existe ainsi deux modèles principaux, soit à partir d'une conception **statique** de la voyelle, représentée par son état stable uniquement constituée des valeurs des trois premiers formants, soit à partir d'une conception plus **dynamique** des voyelles, intégrant des aspects de variations **intrinsèques** de la voyelle (évolution des trajectoires des voyelles pour la distinction entre les voyelles **Tense** vs **Lax**, par exemple, voir Slifka, 2003²) et/ou des aspects de variations **extrinsèques** (concernant les phénomènes de coarticulation consonne-voyelle, représentés par exemple par les transitions formantiques). Winifred Strange (1989) résume ce débat en présentant deux approches perceptives différentes : l'approche « *Simple Target* » où seule la cible vocalique est prise en compte et les approches « *Elaborated Target* » et « *Dynamic Specification* » où les indices **dynamiques** (**intrinsèques** et **extrinsèques**) des voyelles sont pris en compte. Différents auteurs ont montré que la prise en compte des aspects acoustiques **dynamiques** contribuent à une meilleure catégorisation~discrimination du timbre vocalique, mais le débat continue de susciter des controverses (voir Strange, 1989).

Nous avons décidé d'explorer la question du rôle des aspects **dynamiques** en production et en perception des voyelles en arabe. Le choix a porté sur deux dialectes arabes

² Slifka (2003) a montré que la direction de la **pente** de F_1 , qui était descendante pour les voyelles tendues et montante pour les relâchées dans une syllabe de type CVC, permet de discriminer ces deux catégories de voyelles avec un taux de classification correcte proche des 90%.

qui diffèrent en nombre et en qualité vocalique : l'arabe marocain avec trois voyelles pleines « longues » /i: a: u:/ et zéro, une ou deux voyelles brèves (différentes interprétations sont proposées dans la littérature sur le système vocalique de l'arabe marocain : aucune voyelle brève phonologique, un /ə/ ou deux voyelles /ə u/) et l'arabe jordanien avec un système vocalique à cinq voyelles longues /i: e: a: o: u:/ et trois voyelles brèves /i a u/.

Nous nous sommes également placé dans une perspective translinguistique afin de comparer le rôle des aspects **dynamiques** en production et en perception des voyelles de l'arabe en comparaison avec les voyelles du français. Il existe en effet des différences importantes entre les deux langues en question, en termes de nombre de voyelles (au maximum cinq pour l'arabe marocain et huit pour l'arabe jordanien, contre onze voyelles (orales) pour le français) et surtout de la structure morphologique (que nous avons exposée plus haut).

Nous nous attendons à observer des différences entre les voyelles produites et perçues dans les deux langues en ce qui concerne l'importance des indices : les indices **dynamiques** faciliteraient plus la perception des voyelles dans les deux dialectes arabes comparés à la perception des voyelles du français, ces différences étant potentiellement liées au statut et au rôle différents des voyelles entre les deux langues. Afin de vérifier cette hypothèse, nous étudierons tout d'abord le rôle de ces traits **dynamiques** en production des voyelles par le biais de paramètres **dynamiques extrinsèques** (les **pentés formantiques** obtenues par régressions linéaires et polynomiales). Nous comparerons ces mesures **dynamiques** avec les mesures « traditionnelles » des voyelles n'exploitant que les valeurs formantiques obtenues à l'état stable de la voyelle. Nous caractériserons les voyelles en fonction de deux facteurs : l'influence du lieu d'articulation (bilabial, dental et vélaire) et l'influence de la situation de production (en Mot, en Syllabe et en Isolation). En perception, nous évaluerons le rôle des indices **dynamiques** potentiels pour ce qui est de la caractérisation des **prototypes** vocaliques (via une méthode d'ajustement des formants) et de la catégorisation de l'espace vocalique, en comparaison avec les indices **statiques** potentiels.

Cette thèse sera divisée en deux sections : une partie théorique (chapitres 1, 2 & 3) et une partie expérimentale (chapitres 4 & 5). Dans la première partie, une revue de la littérature sera proposée. Dans le Chapitre 1, nous commencerons par expliquer les différences existant entre l'arabe et le français tant au niveau des systèmes vocaliques que morphologiques. Nous nous intéresserons ensuite à la définition de la voyelle en arabe, en fonction des propositions des « grammairiens arabes » et de travaux plus récents. Nous développerons ensuite la situation linguistique de la langue arabe, en abordant son histoire, sa structuration (i.e., différents registres linguistiques) et les différences dialectales existantes (au niveau géolinguistique et sociolinguistique). Nous terminerons ce chapitre par une description approfondie des deux dialectes arabes étudiés dans cette thèse : l'arabe marocain et l'arabe jordanien.

Nous continuerons cette partie théorique en abordant, dans le Chapitre 2, la question de la nature des variations affectant les voyelles produites et de leurs rôles dans la perception des voyelles. Ce sont en effet les différentes caractéristiques liées principalement au phénomène de coarticulation et leurs traitements en production et en perception de la parole qui nous intéressent. Nous étudierons les différences entre les caractéristiques **statiques** et **dynamiques** en fonction de différents modèles de perception des voyelles : les modèles « *Simple Target* », « *Elaborated Target* » et « *Dynamic Specification* », en expliquant les motivations et les justifications pour chaque modèle en faveur des indices soutenus. Les différents indices proposés par ces modèles de perception des voyelles peuvent aller dans le sens d'une théorie active en perception de la parole. Aussi, présenterons-nous différents modèles de perception de la parole, en fonction des modèles passifs (basés sur l'audition) et actifs (basés sur la production), (MacKay, 1956*).

Basée sur ces éléments théoriques, nous proposerons en conclusion de cette section, dans le Chapitre 3, une reformulation de notre problématique et de nos hypothèses de travail, en nous intéressant plus particulièrement aux effets de la densité des systèmes vocaliques sur la taille des espaces vocaliques dans les trois langues. Nous parcourons les différentes théories tentant d'expliquer les tendances observées dans les langues (i.e., la **théorie quantique** proposée par Kenneth Stevens (Stevens, 1972, 1989, 2003) ; la **théorie de la**

dispersion adaptée proposé par Björn Lindblom (Lindblom, 1986) et la **théorie de dispersion-focalisation**, proposée par Schwartz *et al.*, 1989 ; Vallée, 1994 ; Boë *et al.*, 1996 ; Schwartz *et al.*, 1997_a, 1997_b).

Dans la deuxième partie, nous présenterons le cadre expérimental de ce travail. Nous nous intéresserons dans le Chapitre 4 aux résultats en production des voyelles. En introduction, nous donnerons une description de la méthodologie du travail en nous intéressant plus particulièrement aux locuteurs, aux corpus, aux protocoles expérimentaux et aux méthodes d'analyse. Ce chapitre sera divisé en deux sous-parties. Dans la première sous-partie, nous commencerons par présenter les résultats de la comparaison inter-langues, en développant tout d'abord les résultats obtenus en fonction de caractéristiques **statiques** et ensuite, nous présenterons les résultats en fonction de différentes analyses **dynamiques**. Nous terminerons par la comparaison des différents résultats **statiques** et **dynamiques** obtenus en évaluant leur pertinence par une analyse discriminante. Dans la deuxième sous-partie, nous présenterons les spécificités de chaque langue en fonction des résultats **statiques** et **dynamiques**, en nous intéressant plus particulièrement aux voyelles problématiques dans chacune des langues : les voyelles brèves en arabe marocain et arabe jordanien ; et les voyelles dites « médianes » et le statut des deux voyelles ouvertes /a a/ en français. À chaque fois, nous évaluerons la validité des résultats obtenus par une analyse discriminante.

Nous nous intéresserons dans le cinquième et dernier chapitre aux résultats en perception des voyelles. En guise d'introduction de ce chapitre, nous présenterons les locuteurs, les corpus utilisés, les protocoles expérimentaux et les analyses proposées. Ce chapitre sera divisé en deux sous-parties. Nous commencerons par présenter les résultats obtenus à partir d'une première expérience sur l'identification de **prototypes** vocaliques (via une méthode d'ajustement des formants), dans chaque langue et ensuite dans une comparaison inter-langues en fonction des paramètres utilisés, **statiques** (*V Target*) et **dynamiques** (*CV Target*). Là encore, nous vérifierons la pertinence de ces paramètres par une analyse discriminante. Nous terminerons par la présentation des résultats obtenus à partir de la seconde expérience de perception des voyelles sur la catégorisation des espaces vocaliques,

dans les trois langues en fonction des paramètres utilisés, **statiques** (*V Target*) et **dynamiques** (*CV Target*).

Nous terminerons par une conclusion générale et les perspectives issues de ce travail, en insistant sur les questions auxquelles nous n'avons pas pu répondre de manière définitive.

Section I. CADRE THÉORIQUE

Chapitre 1.

LES SYSTÈMES VOCALIQUES DE L'ARABE

*« All Semitic languages – Hebrew, Arabic, Phoenician and the rest – possess in common a peculiar devotion to consonants. In fact, a Semite does not think of a Semitic word as being composed of syllables, he thinks of it as being made of the strong bones of consonants with the vowel sounds floating above like invisible spirits. Moreover, the vowels of a Hebrew or Arabic word have little to do with the determination of meaning. Meaning is firmly staked out by the consonants alone... » (nous soulignons).
(Burgess, 1992 : p. 78).*

1.1. INTRODUCTION

Qu'est-ce que la voyelle en arabe et comment les sujets arabophones la produisent et la perçoivent ? Ce sont quelques questionnements qui nous préoccupent et auxquelles nous tenterons de répondre dans cette thèse.

Les sujets arabophones jordaniens et marocains ont montré, dans nos études antérieures, de grandes difficultés à produire et à percevoir les voyelles en isolation (voir Al-Tamimi, 2002 ; Al-Tamimi & Barkat-Defradas, 2002 ; Barkat-Defradas *et al.*, 2003). Par exemple, certains sujets tendaient à produire toutes les voyelles isolées et quelque soit la cible, proches soit de /a:/ soit de /ə/. En perception, à partir d'une méthode d'ajustement de formants (Johnson *et al.*, 1993), les locuteurs arabophones identifiaient les prototypes vocaliques de leur système. Ces **prototypes** vocaliques obtenus se situaient dans les extrémités de l'espace vocalique avec une grande variabilité inter-individuelle. Les locuteurs arabophones nous ont exprimé les grandes difficultés qu'ils avaient à comprendre la tâche

demandée dans les deux expériences, et plus particulièrement en perception : ils avaient tendances à entendre des « bips » ou des « bruits ».

Nous avons pensé que ces mêmes difficultés devaient affecter les réponses des auditeurs français ou italiens. Or, ces derniers arrivaient plus facilement à produire et à percevoir les voyelles en isolation (voir Girard, 2002 ; Carré & Hombert, 2002).

Ces différences de comportements entre les locuteurs arabophones d'un côté et les locuteurs français et italiens de l'autre nous ont conduit à supposer que l'origine de ces difficultés est intimement liée au statut des voyelles dans ces langues. Nous avons formulé alors deux hypothèses tendant à expliquer ces différences par : a) la morphologie du lexique qui est à l'origine d'un statut différent des voyelles et b) l'utilisation d'indices statiques (i.e., cibles vocaliques)³.

Sur le plan morphologique, on sait que l'organisation du lexique dans toutes les langues sémitiques, dont l'arabe, est dépendante des consonnes, qui ont un rôle important pour véhiculer le sens des mots. Les voyelles peuvent être considérées alors comme des outils grammaticaux aidant à la construction des mots (voir Versteegh, 2001 ; Ryding, 2005). La situation des voyelles en français et dans les langues indo-européennes est complètement différente.

La langue arabe est une langue non-concaténative à racine consonantique trilitère qui exhibe des relations consonnes~consonnes directes pour structurer son lexique (McCarthy, 1979 ; Pierrehumbert, 1992, etc...). Les différents auteurs expliquent que certaines catégories verbales, comme l'aspect ou la voix, sont marquées par la modification systématique de la-(des) qualité(s) vocalique(s) du schème (ou du **pattern vocalique**, voir *infra*), sans aucune modification de la **racine de base** (voir McCarthy, 1979, 1981, 1982 ; McCarthy & Prince, 1988, 1990, etc...). Le français quant à lui présente une structure à dominance concaténative basée sur une morphologie flexionnelle.

³ Nous reviendrons sur cette question dans le chapitre 2.

La littérature explique qu'en arabe **classique** un mot est composé de deux parties distinctes : une **racine de base**, généralement trilitère, et un **pattern vocalique** constitué soit d'une ou plusieurs voyelles, soit de consonnes auxiliaires⁴, soit des deux ensembles (seuls ou en combinaison avec des affixes dérivationnels). La différence entre ces deux parties est que la première véhicule le sens lexical d'un mot, la seconde le sens grammatical⁵.

Roman (1981/82_a) explique pourquoi les **racines de base** en arabe ne peuvent être que consonantiques en donnant les trois raisons indissociables suivantes (Roman, 1981/82_a : pp. 145) :

- Toute voyelle radicale longue ne pourrait appartenir à une syllabe CVC parce que celle-ci aurait trois mores et serait anormale⁶,
- Aucune forme ne peut commencer par un élément radical puisqu'il n'existe pas de syllabe VC (ou implosive) avec laquelle des hiatus seraient réalisés dans les séquences de type : CV (explosive) + VC (implosive), et enfin
- La langue arabe étant sans hiatus, les voyelles ne peuvent donc prendre place dans une forme au contact d'une radicale qui serait une voyelle.

Roman (1981/82_a) explique également que l'arabe **classique** a gardé du protosémitique un système syllabique simple constitué des formes /CV/ et /CVC/.

Cette division du mot en arabe (et dans les langues sémitiques) en **racine de base** et en **pattern vocalique** peut être mieux cernée en utilisant les nouvelles théories phonologiques non-linéaires : en Phonologie Autosegmentale et plus particulièrement en Phonologie

⁴ Ces consonnes auxiliaires sont : /ʔ w j m t n h s l/.

⁵ Pour plus de détails sur la structure de l'arabe **classique**, voir entre autres : Versteegh (2001, Chapitre 6) et Ryding (2005).

⁶ Roman (1981/82_a) explique que les unités minimales prononçables en arabe seraient constituées des trois unités métriques fondamentales : brève, longue et surlongue, qui valent, respectivement, une more (v), deux mores (vv) et trois mores (vvv). L'unité brève apparaît toujours composée d'une consonne et d'une voyelle, l'unité longue, soit d'une consonne et d'une voyelle longue, soit d'une consonne, d'une voyelle brève et d'une consonne, et l'unité surlongue soit d'une consonne, d'une voyelle longue et d'une consonne, soit d'une consonne, d'une voyelle brève ou longue et de deux consonnes. Les unités surlongues (vvv) sont, d'après l'auteur, des variantes conditionnées phonétiquement ou syntaxiquement, et sont donc considérées comme anormales et non présentes comme unité canonique de la koinè arabe (Roman, 1981/82_a : pp. 136-137).

Squelettale (Goldsmith, 1976_a, 1976_b & 1990 ; Kenstowicz, 1994) et en Morphologie de Gabarits ou Morphologie Prosodique (McCarthy & Prince, 1990)⁷. Pour ces différentes théories, la séparation des consonnes et des voyelles en **racine de base** et en **patterns vocaliques** facilite l'étude de la morphologie de l'arabe (et des langues sémitiques) car elle permet une meilleure appréhension du rôle des consonnes et des voyelles dans la construction du lexique (voir l'exemple dans la Figure 1).

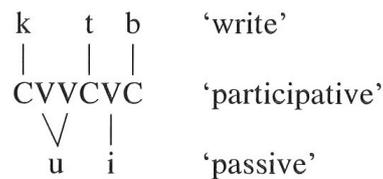


Figure 1 : Représentation du mot /kutib/ « On a correspondu avec » en Phonologie Squelettale. (Emprunté à Kenstowicz, 1994).

Dans cet exemple, la **racine de base K.T.B** « écrire » est utilisée pour former le passif de la troisième forme radicale de l'arabe /ka:taba/ « correspondre » : /kutib/ signifiant « on a correspondu avec » en utilisant la forme participative, avec la structure CVVCVC et le **pattern vocalique** indiquant le passif /u:~i/.

La grande différence entre la morphologie du lexique en français et en arabe peut peut-être expliquer les différences quant au statut et au rôle relatifs des voyelles dans la morphologie du lexique des deux langues. Les contraintes phonotactiques en arabe et en français sont différentes : la voyelle en arabe ne peut être produite qu'en association avec une consonne (voir *infra*), en français, elle peut être présente à l'initiale des mots et même en isolation.

Outre les différences sur le plan morphologique entre l'arabe et le français, nous observons des différences liées aux systèmes vocaliques mêmes : l'arabe marocain présente un système vocalique composé de cinq voyelles /i ə a: u u:/ ; l'arabe jordanien, un système à huit voyelles /i i: e: a a: o: u u:/ et le français présente un système vocalique dense composé

⁷ Pour d'autres approches nouvelles en morphologie du lexique de la langue arabe, voir les travaux de Georges Bohas (e.g., Bohas, 1981, 1982, 1985, 1993 ; Ayoub & Bohas, 1983 ; Bohas & Guillaume, 1984, Bohas *et al.*, 1990, etc...).

de onze voyelles orales /i e ε a α ɔ o u y ø œ/, tandis que et. En plus de la différence en nombre de voyelles, le système vocalique de l'arabe marocain a tendance à se centraliser par rapport à celui de l'arabe jordanien (ces effets sont observés entre l'arabe maghrébin et l'arabe moyen-oriental en général, voir Barkat, 2000 ; et entre les deux dialectes, voir Al-Tamimi & Barkat-Defradas, 2002 ; Barkat-Defradas *et al.*, 2003; Al-Tamimi *et al.*, 2004; Al-Tamimi & Ferragne, 2005 ; Al-Tamimi, 2006 ; Al-Tamimi, 2007). Les locuteurs français tendaient à produire des voyelles positionnées dans les extrémités de l'espace vocalique : le nombre important de segments vocaliques en français demande une certaine précision articulatoire et une distance perceptive suffisante permettant de distinguer ses voyelles.

Étant donné les différences phonotactiques entre l'arabe et le français, les voyelles en arabe ont un statut différent. Afin de comprendre cette question, nous nous intéressons dans la partie qui suit à la notion de la voyelle en arabe, à travers les descriptions des grammairiens arabes, mais également au niveau de l'apport des études dialectologiques actuelles à sa caractérisation. Nous allons nous intéresser dans notre travail aux systèmes vocaliques de deux dialectes arabes : l'arabe marocain et l'arabe jordanien. Auparavant, nous situerons ces deux dialectes dans une présentation des caractéristiques générales de la langue arabe en commençant par son histoire, sa structuration et les classifications de ses dialectes modernes.

1.2. LA VOYELLE EN ARABE

La grande majorité des linguistes arabisants ont étudié le système vocalique de l'arabe **classique** et ont retenu un système constitué de trois timbres avec une opposition de durée (trois segments longs et trois segments brefs), (voir D. Cohen, 1962, 1972, 1988, etc...). Cantineau (1960) explique que le sémitique et l'arabe **classique** (et ancien) présentaient un système triangulaire à trois timbres vocalique /i a u/ avec une opposition de durée.

En travaillant dans le domaine de la dialectologie arabe, nous pouvons nous poser les questions suivantes : quelle est la définition de la voyelle en arabe ? Quel est son statut au sein de la morphologie du lexique ? Qu'apportent les différentes études phonético-phonologiques actuelles dans la compréhension des systèmes vocaliques en arabe (dialectal et standard

contemporain) ? Commençons tout d'abord par comprendre la notion de la voyelle en arabe en partant des descriptions de quelques grammairiens arabes et en terminant par la présentation de quelques études qui se sont intéressées à la voyelle en arabe

1.2.1. *Vision des grammairiens arabes*

Différents grammairiens arabes ont tenté de définir la voyelle en arabe. C'est le cas de Sibawayhi dans *Al-Kitāb* (Sibawayhi, 8^{ème} siècle, également dans Haroun, 1983), Ibn Jinni dans *Sir Šinā'at Al-ʿi'rāb* (Ibn Jinni, 10^{ème} siècle, également dans Mehiri, 1973) et Cantineau (1960) (qui a plus ou moins repris les travaux des anciens grammairiens arabes), pour n'en citer que quelques uns.

Pour Sibawayhi, la voyelle constitue un « outil supplémentaire qui sert à produire les consonnes. Les voyelles sont divisées en deux catégories : les lettres de prolongation (ou les voyelles longues) et les lettres de mouvement (de la consonne) (ou les voyelles brèves) ». La lettre chez Sibawayhi est constituée de deux parties : une partie stable et une partie en mouvement. Il définit la lettre comme « une unité sonore de la parole composée d'une partie stable (ou consonne) et d'une partie supplémentaire (une voyelle), et c'est cette partie supplémentaire qui détermine si la lettre est en mouvement ou non ». Sibawayhi a défini trois timbres vocaliques en arabe, /i a u/ qui s'ajoute aux consonnes pour les produire. Pour lui, puisqu'une lettre de prolongation est composée d'un timbre vocalique (en mouvement) et d'une consonne, toutes les lettres de prolongation sont des lettres « akinésiques » qui n'ont pas de mouvement d'articulateurs (voir *infra* la définition proposée par Cantineau, 1960). On peut alors dire qu'une lettre « kinésique » (qui se produit avec un mouvement des articulateurs) ou une « lettre en mouvement » est composée d'une partie « akinésique » (une consonne) et d'un mouvement, ce qui donne une production explosive, i.e., un « CV ». Tandis que pour une lettre « akinésique », c'est la partie sonore qui se prononce en fermeture, i.e., « VC ». (vision développée également dans Braham, 1997).

Pour Ibn Jinni, le système vocalique de l'arabe est identique à celui du Sémitique, à savoir un système à trois timbres vocaliques (trois segments longs et trois segments brefs). Il a

été le premier à définir la voyelle en arabe clairement et à consacrer un chapitre entier à la phonétique et à la phonologie de l'arabe. Pour lui, les voyelles longues constituent les lettres de prolongation, tandis que les voyelles brèves constituent une sous-classe de celles-ci. Comme ses prédécesseurs, Ibn Jinni s'est intéressé aux variations allophoniques des timbres vocaliques en arabe. L'auteur explique différents processus d'assimilation qui, selon lui, augmentent le nombre de timbres vocaliques de trois à six (trois timbres vocaliques initiaux (constitués de trois voyelles longues /i: a: u:/ et trois brèves /i a u/) et trois timbres intermédiaires (des allophones)), (voir Cantineau, 1960 : pp. 96-102, pour des exemples de ces trois phénomènes) :

- L'imāla⁸ qui se produit quand le /a/ s'assimile vers un /i/. Ici on a deux cas :
 1. Une imāla forte, ce qui donne un [e:],
 2. Une imāla légère, ce qui donne un [ɛ:],
- L'emphatisation qui se produit quand il s'agit d'une production de /a:/ en contexte emphatique, ce qui donne un [ɑ:],
- L'išmām qui se produit quand /i:/ est incliné vers /u:/, donnant ainsi [y:].

Comme Sibawayhi, Ibn Jinni considère la voyelle comme faisant partie intégrante de la consonne et elle n'est qu'un outil facilitant leur production. En effet, tous les grammairiens arabes expliquent que la voyelle n'existe pas en isolation : elle est présente en association (et même en appartenance) à la consonne, résultant ainsi par la non-existence des séquences V, VCV, etc... en arabe (Cantineau (1960), etc...). Le plus souvent, on observe la réalisation d'un coup de glotte avant ces séquences, i.e., V > ʔV, VCV > ʔVCV.

Jean Cantineau (1960) propose, dans son livre « Étude de Linguistique Arabe », une description des systèmes vocaliques et consonantiques de l'arabe classique et de l'arabe dialectal (de différentes zones géographiques). En se basant sur les travaux de différents

⁸ L'imāla correspond à une antériorisation et une fermeture de la voyelle /a/ selon les contextes, allant jusqu'à une réalisation de /a/ en /i/. Dans certains parlers du Moyen-Orient la réalisation extra-forte de l'imāla, où le /a/ est totalement assimilé à un /i/

grammairiens arabes, l'auteur décrit le système vocalique de l'arabe comme ses prédécesseurs (i.e., trois longs et trois brefs).

Il définit la voyelle, en disant que : « les grammairiens arabes distinguent trois timbres vocaliques : fath = a, kasr = i, dam = u, avec diverses nuances secondaires. Les voyelles longues sont considérées comme les résultantes de la combinaison d'une voyelle brève (lettre de mouvement) avec une des semi-voyelles w, y ou alif. Ces voyelles longues sont appelées ainsi les lettres de prolongation. Elles sont alors réalisées comme des diphtongues : /u:/ ⇒ uw, /i:/ ⇒ ij, /a:/ ⇒ a + alif ».

Outre les assimilations augmentant le nombre de timbres vocaliques en arabe, Cantineau explique la présence d'autres phénomènes produisant soit un allongement soit une réduction de la durée vocalique. L'allongement vocalique peut être produit dans la récitation du Coran, par exemple, où certaines voyelles longues deviennent ultra-longues (de plusieurs secondes !).

Pour ce qui concerne la réduction des voyelles, Cantineau introduit une nouvelle catégorie vocalique : les voyelles ultra-brèves, présentes dès l'arabe ancien. Il explique la présence de cette « nouvelle » voyelle en donnant l'exemple de la structure syllabique suivante : CVCCVCVV. Cette structure syllabique devient en arabe CVCCCVV par suite de chute de la voyelle brève en syllabe ouverte. Mais si la suite des trois consonnes est difficilement prononçable, une voyelle de disjonction (d'abord ultra-brève, pouvant devenir par la suite brève) va s'ajouter et la structure syllabique devient CVCVCCVV (ex. *nidrisuu ⇒ *nedrsuu ⇒ nedərsuu « nous étudions »).

Cantineau explique ensuite les variations existantes dans les systèmes vocaliques des différents dialectes arabes. Ses descriptions vont dans le sens de la plupart des études actuelles en dialectologie, à savoir le maintien des trois voyelles « longues » ou pleines dans tous les dialectes arabes, le passage des voyelles /aj aw/ soit à des voyelles périphériques /e: o:/ dans les dialectes moyen-orientaux, soit à des voyelles fermées /i: u:/ dans les dialectes maghrébins, soit leur maintien sans modification. Quant aux anciennes voyelles brèves, Cantineau (1960) explique le maintien de celles-ci dans les dialectes moyen-orientaux et la

présence d'opposition binaire seulement dans les dialectes maghrébins (/ə u/, où les anciennes voyelles /i a/ ont fusionné pour devenir /ə/, voir *infra*).

Cette séparation des rôles des consonnes et des voyelles va dans le sens de ce que nous ont appris les grammairiens arabes, i.e., la voyelle ne peut exister qu'en présence et même en appartenance à la consonne. Nous pouvons alors nous poser la question de savoir comment la voyelle est produite et surtout perçue par les locuteurs~auditeurs natifs ? Est-ce que les voyelles n'ont aucun rôle phonologique dans la langue et par conséquent leur perception posera des problèmes aux arabophones ? Que représentent-elles pour les auditeurs sur le plan cognitif ? Quel est l'apport des consonnes dans la reconnaissance des voyelles en arabe ? Voici quelques-unes des questions auxquelles nous tenterons de répondre dans cette thèse.

Nous nous intéressons dans la partie qui suit à quelques travaux qui ont exploré les systèmes vocaliques de l'arabe standard contemporain et de l'arabe dialectal du point de vue acoustique et perceptif.

1.2.2. *Les études actuelles*⁹

Plusieurs linguistes se sont consacrés à l'étude phonologique des systèmes vocaliques et consonantiques de l'arabe standard moderne et de l'arabe dialectal. Nous commençons par Al-Ani (1970) qui a étudié le système vocalique de l'arabe standard réalisé par des locuteurs irakiens (neuf locuteurs et une locutrice). L'auteur a demandé aux locuteurs de produire les voyelles longues et brèves de l'arabe standard en isolation. Les résultats qu'il a obtenus montrent que, pour toutes les catégories vocaliques confondues, les voyelles brèves ont une durée de 300 ms, comparées aux longues de durée de 600 ms. Il a observé également des timbres vocaliques identiques entre les voyelles longues et brèves. Pour Al-Ani (1970), les systèmes vocaliques de l'arabe standard et de l'arabe dialectal, sont identiques, présentant trois

⁹ Beaucoup d'études sur le système vocalique de l'arabe standard moderne et/ou l'arabe dialectal se sont concentrées sur l'analyse de la durée vocalique et les rapports entre voyelles longues et voyelles brèves (voir les différents travaux de Mounir Jomaa, de Laura Abou-Haidar ou de Mohammed Embarki). Ce qui nous intéresse dans cette partie sont les quelques études sur les qualités mêmes des systèmes vocaliques de l'arabe.

voyelles longues et trois voyelles brèves sans distinction de timbre mais avec une durée vocalique deux fois plus longues pour les voyelles longues par rapport aux voyelles brèves.

Le problème posé par les résultats de l'étude d'Al-Ani réside dans la réalisation non naturelle des voyelles à la fois en isolation et hors contexte et à débit lent. C'est ce qu'a reproché Ghazali (1979) à Al-Ani. L'étude proposée par Salem Ghazali constitue la première étude trans-dialectale effectuée sur l'arabe dialectal et non l'arabe standard. Douze locuteurs de six dialectes arabes différents (deux locuteurs par dialecte, les dialectes sont l'arabe algérien, tunisien, libyen, égyptien, jordanien et iraquien) ont participé à ce travail, qui avait pour but l'étude des variations de la durée vocalique et de ses effets sur les caractéristiques formantiques des voyelles. Il est le premier à rejeter l'approche de Cantineau et d'Al-Ani et à avancer l'idée selon laquelle le système vocalique de l'arabe n'est pas constitué de trois timbres vocaliques avec une opposition de durée comme on le dit traditionnellement, mais d'un système à trois voyelles longues et un autre à trois voyelles brèves (ces voyelles brèves diffèrent des longues en durée et en timbre vocalique). Il a aussi étudié les effets de la pharyngalisation sur les voyelles. Les résultats obtenus vont dans le sens d'une distinction au niveau qualitatif et quantitatif des timbres vocaliques de l'arabe dialectal. Néanmoins, la notation qu'il a utilisée ne montre pas cette différence, car pour lui le système vocalique phonologique de l'arabe dialectal correspond à /i i: a a: u u:/, la différenciation des timbres se trouve au niveau phonétique.

La plupart des études actuelles ont porté sur l'analyse du système vocalique de l'arabe standard. Nous nous référons essentiellement à l'étude proposée par Belkaid (1984). L'auteur a étudié le système vocalique de l'arabe standard basé sur un seul locuteur d'origine tunisienne. Son corpus est basé sur 70 mots (50 de deux syllabes et 20 de trois) comportant 110 voyelles brèves et 50 voyelles longues, la voyelle /a/ étant présente dans son corpus à hauteur de 50% des voyelles brèves et de 55% des voyelles longues. L'auteur a trouvé qu'il existe un changement de timbre en fonction de la durée des segments, i.e., six timbres différents avec des durées variables existent en arabe standard au lieu des trois timbres dédoublés en fonction de la durée (voir également les résultats sur les changements de timbre vocalique en fonction de la durée et de l'accent pour /a a:/, obtenus par de Jong & Zawaydeh, 2002).

Les études proposées par Imad Znagui (Znagui, 1992, 1993, 1995 ; Znagui & Boudelaa, 1993 ; Znagui & Bouchnafi, 1994 ; Znagui & Yeou, 1996) ont porté sur la production et la perception des voyelles de l'arabe standard, par six locuteurs d'origine maghrébine. Les résultats obtenus, en production de la parole, montrent l'existence d'une influence du contexte environnant sur les timbres vocaliques au niveau phonétique uniquement. L'auteur a divisé le système vocalique en deux groupes : les voyelles antériorisées [i i: æ æ: u u:] en présence de consonnes d'avant et les voyelles postériorisées [e e: a a: o o:] en présence de consonnes d'arrières et de consonnes pharyngalisées. Pour trouver une corrélation entre production et perception de la parole, i.e., pour défendre les propos avancés par la Théorie Motrice de la parole (Lieberman *et al.*, 1967_a) (voir § 2.3.2.1), l'auteur a réalisé des expériences perceptives lui permettant de tester la validité de cette théorie. Les résultats obtenus en perception de la parole montrent que les auditeurs ne sont pas sensibles aux variations phonétiques obtenues en production de la parole, validant ainsi les théories auditives de la perception de la parole selon laquelle les auditeurs ne se réfèrent pas à leurs connaissances articulatoires pour percevoir les sons de leurs langues maternelles.

L'auteur avance l'idée selon laquelle les informations **dynamiques** contenues dans la coarticulation CV (i.e., les transitions formantiques) et les distances entre les deux premiers formants et non leurs valeurs absolues, doivent être prises en compte pour l'analyse des voyelles de l'arabe standard. D'autre part, l'auteur a cherché à valider la notion de voyelles tendues *vs* relâchées en arabe (comme en anglais) dépendant de l'accent du mot. Ses résultats montrent qu'il n'y a pas d'influence de l'accent sur la dispersion des voyelles en arabe standard. Il en déduit donc que le système vocalique de l'arabe ne peut entretenir une différence fonctionnelle au niveau de l'accent et de la durée à la fois (donc pas de distinction voyelles tendues *vs* voyelles relâchées en arabe, contrairement aux résultats obtenus par Shahin, 1997, 2003, qui a montré que le système vocalique de l'arabe palestinien bénéficie de cette distinction entre voyelles longues et brèves, et qu'il existe également une opposition de type [±ATR]). (Voir également les travaux de Herzallah, 1990* ; Younes, 1982*, 1984).

Une question peut émerger alors : existe-il vraiment des différences en production et en perception de la parole entre les systèmes vocaliques de l'arabe standard et de l'arabe

dialectal ? Comment peut-on évaluer cette distinction de type voyelles tendues vs voyelles relâchées en production et en perception de la parole ?

Dans une précédente étude (Al-Tamimi, 2001), nous avons réalisé une description phonético-phonologique de six parlers jordano-palestiniens. Nous avons étudié les interférences consonantiques et vocaliques entre arabe standard et dialectal pour différents locuteurs jordaniens. Nous avons élaboré un corpus ayant des items lexicaux identiques en arabe standard et dialectal, construit avec différentes structures syllabiques et différentes consonnes potentiellement révélatrices [s s^f t t^f d d^f θ ð ð^f q k ʒ] (i.e., utilisées largement pour l'opposition arabe standard vs arabe dialectal). Après avoir vérifié que le système vocalique (et consonantique) de ces différents locuteurs était homogène (i.e., système vocalique à huit voyelles /i iː a aː u uː e oː/), nous avons trouvé que l'espace de production des voyelles de l'arabe dialectal est différent de celui des voyelles de l'arabe standard (produits par les mêmes locuteurs et pour les mêmes mots) : l'espace vocalique en arabe standard est significativement plus postérieur que celui en arabe dialectal¹⁰.

Par ailleurs, Barkat (2000) a effectué une étude trans-dialectale au niveau des systèmes vocaliques des dialectes arabes. Dans sa thèse, Mélissa Barkat a établi une typologie des indices acoustiques robustes pour permettre l'identification automatique des dialectes arabes. Vingt-quatre locuteurs (12 hommes et 12 femmes) originaires de deux zones dialectales maghrébine et moyen-orientale (voir § 1.3.3.1) (six dialectes arabes : l'arabe marocain, algérien, tunisien, jordanien, libanais et syrien) ont produit le texte « la bise et le soleil » semi-spontanément¹¹. Au total, le corpus contenait près de 1500 voyelles toutes positions accentuelles, structures syllabiques, ou contextes consonantiques confondus. Afin d'homogénéiser les analyses entreprises en fonction des différences de durée et/ou des contextes syllabiques et consonantiques, les valeurs formantiques ont été relevées au milieu

¹⁰ Vu les différences observées en production de la parole, nous nous interrogeons sur les possibles différences existantes au niveau perceptif ? Un thème de recherche qu'il serait judicieux de développer, car apportant un autre paramètre de discrimination entre l'arabe standard et l'arabe dialectal.

¹¹ Chaque phrase a été produite par l'expérimentatrice à voix haute dans une langue autre que l'arabe (i.e., soit le français soit l'anglais) et les locuteurs produisaient la phrase qu'ils avaient comprise quatre fois de suite en arabe dialectal.

temporel des voyelles. L'auteur a comparé les réalisations phonétiques vocaliques de différents locuteurs et a établi une typologie des dialectes arabes basée sur l'opposition : Maghreb vs Moyen-Orient et en fonction du nombre de segments vocaliques. Un des principaux résultats est que l'espace vocalique des dialectes arabes du Maghreb est plus centralisé que celui des dialectes moyen-orientaux. Il existe également une différence qualitative et quantitative entre les voyelles longues et brèves dans tous les dialectes, que nous présentons avec une notation phonologique (l'analyse établie par Barkat, 2000, ne mentionne pas si les systèmes vocaliques doivent être considérés du point de vue phonétique ou phonologique) :

- Le dialecte Algérien d'Alger, avec cinq timbres vocaliques (six voyelles) :
/i: ə æ æ: a: u:/.
- Le dialecte Marocain de Casablanca, avec quatre timbres vocaliques (cinq voyelles) :
/i: ə a: u u:/.
- Le dialecte Tunisien de Tunis avec cinq timbres vocaliques en opposition de durée (neuf voyelles) : /i i: ə ə: æ æ: a a: u u:/.
- Le dialecte Jordanien d'Irbid, avec cinq timbres vocaliques (huit voyelles) :
/i i: e: a a: o: u u:/ (les deux voyelles /e: o:/ représentent les anciennes diphtongues /aj aw/ qui sont soit réduites, soit maintenues dans certains mots).
- Le dialecte Libanais de Beyrouth, avec cinq timbres vocaliques (six voyelles et deux diphtongues) : /i: e e: a o u: aj aw/.
- Le dialecte Syrien de Damas, avec cinq timbres vocaliques (huit voyelles) :
/i i: e: a a: o: u u:/).

L'étude de Barkat (2000) fut la première en son genre à s'occuper à la fois de l'analyse des qualités des timbres vocaliques et de leurs oppositions de durée en parole « semi spontanée » dans différents dialectes arabes. La seule limitation de cette étude est la comparaison des voyelles dans des structures syllabiques et des contextes consonantiques différents, ainsi que le faible nombre d'occurrences. On peut se poser la question de savoir si les différentes qualités vocaliques observées dans les dialectes arabes sont dépendantes des structures syllabiques et/ou consonantiques différentes ?

C'est une des raisons qui nous a poussé à établir une étude comparative sur deux dialectes arabes seulement faisant partie des deux zones dialectales du Maghreb et du Moyen-Orient, à savoir l'arabe marocain et l'arabe jordanien, respectivement. Des analyses fines des timbres vocaliques seront proposées, dans des structures syllabiques et des contextes consonantiques contrôlés. Le but est d'étudier l'effet du lieu d'articulation et des contextes de présentation sur les qualités des timbres dans les deux dialectes, en production et en perception de la parole.

Dans une étude antérieure, (Al-Tamimi & Barkat-Defradas, 2002 ; Defradas *et al.*, 2003), les résultats obtenus en production montrent le passage du système vocalique de l'arabe marocain de cinq à quatre timbres vocaliques : /i: ə a: u u:/ > /i: ə a: u:/ en contexte dental /d d^ʕ/. Ces résultats en production ont également été confirmés en perception, où les deux voyelles /ə u/ fusionnent et deviennent /ə/. Étant donnée que ces résultats sont limités au seul contexte dental /d d^ʕ/, nous nous posons la question sur les conséquences de variation des lieux d'articulation des consonnes adjacentes sur la qualité de ces deux voyelles en arabe marocain ? Obtiendrons-nous toujours une seule voyelle brève ou deux ?

Nous venons de présenter la « voyelle » en arabe : sa définition depuis les grammairiens arabes en passant par quelques études actuelles qui ont tenté de la décrire. Nous avons examiné les différences existantes au niveau inter-dialectal et inter zone. Avant de décrire les deux dialectes que nous avons utilisés pour notre étude, nous allons les situer dans le contexte de la langue arabe en général : selon son histoire, sa structure et en proposant une typologie des classifications géolinguistique et sociolinguistique de ses dialectes modernes.

1.3. CARACTÉRISTIQUES GÉNÉRALES DE LA LANGUE ARABE

La dialectologie arabe est un domaine d'étude à la fois riche, intéressant et complexe. Beaucoup de linguistes ont consacré une grande partie de leurs travaux à cette question (voir par exemple W. Marçais, 1902* ; P. Marçais, 1956* ; M. Cohen, 1912*, 1924 ; D. Cohen, 1962, 1965, 1973*, 1988 ; Cantineau, 1955*, 1960 ; Fischer, 1967*, 1982* ; Fischer &

Jastrow, 1980* ; Al-Ani, 1970 ; Ali & Daniloff, 1972* ; Colin, 1975 ; Ghazali, 1977, 1979 ; Ghazali, 1977, 1981 ; Kaye, 2001 ; Versteegh, 2001 ; etc...). Nous rappelons brièvement l'histoire de la langue arabe, sa structure et la situation dialectale actuelle selon les classifications géolinguistique et sociolinguistique.

1.3.1. Histoire de la langue Arabe

La littérature sur l'histoire de la langue arabe est abondante (voir les travaux de Brockelmann, 1910* ; M. Cohen, 1924 ; D. Cohen, 1962, 1968, 1988 ; Cantineau, 1960 ; Roman, 1981/82_b ; Martinet, 1953 ; Hetzron, 1974, 1976, 2001_a, 2001_b ; Bonnot, 1976 ; Ghazali, 1977 ; Bomhard, 1995 ; Hayward, 2000 ; Kaye, 2001 ; Versteegh, 2001 (Chapitre 2) ; etc...).

Les langues faisant partie de la famille **chamito-sémitique** — appelée également **afro-asiatique** — couvrent une aire géographique considérable, qui s'étend du nord de l'Afrique (du Maghreb jusqu'au Nigeria et une partie du Cameroun, en passant par l'Éthiopie, l'Érythrée et la Somalie) et l'île de Malte, ainsi que tout le Proche-Orient, pour s'arrêter aux frontières de l'Iran (quelques îlots d'arabophones). La Figure 2 montre l'aire géographique de la famille **chamito-sémitique**.

L'appellation de **chamito-sémitique** attribuée à ces langues est une invention des linguistiques de la fin du 19^{ème} siècle. Sous l'influence de la Genèse, ceux-ci présentèrent les Hébreux, les Araméens, les anciens Égyptiens et les Arabes comme les descendants de **Sem** (d'où **Sémitique**) et de **Cham** (d'où **Chamite**), les fils du patriarche **Noé**. Quant à **Koush**, un fils de Cham, dont les descendants auraient habité le sud de l'Égypte, il aurait donné son nom à l'Éthiopie, d'où la création par la suite du terme **couchitique** pour désigner les langues de ce pays. On a inventé plus récemment le mot **tchadique** pour désigner les langues de la région du lac Tchad. Enfin, les langues dites **omotiques** sont parlées en Éthiopie dans la région du fleuve Omo.

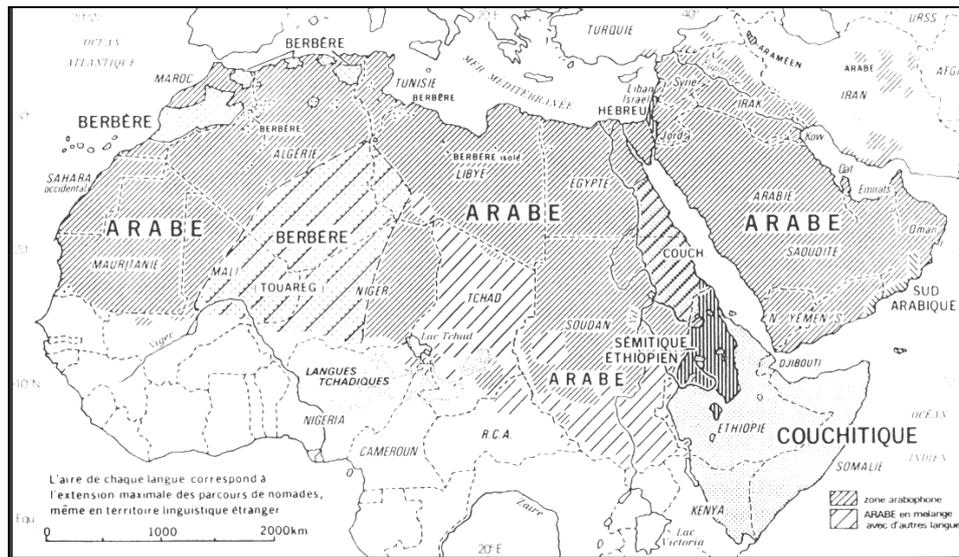


Figure 2 : Distribution géographique des langues **chamito-sémitiques** ou **afro-asiatiques**. (Emprunté à Cohen, 1988).

La Figure 3 montre la classification établie par Hayward (2000) des langues faisant partie du groupe de l’Afro-Asiatique. C’est la classification la plus acceptée actuellement (voir Hayward, 2000, pour plus de détails).

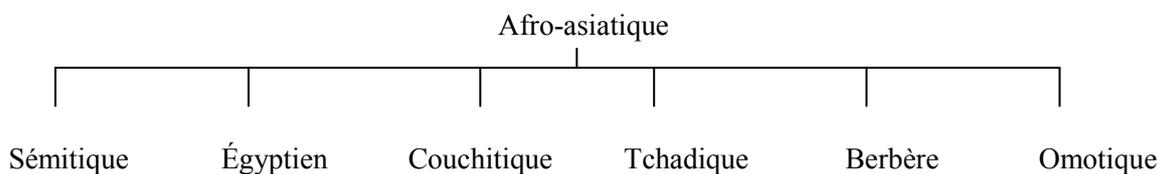


Figure 3 : Classification des langues Afro-Asiatiques. (Adapté de Hayward, 2000).

Certaines des langues de la famille chamito-sémitique ont, dans l’Antiquité, été de très grandes langues de civilisation. Pensons à l'égyptien, au babylonien, au sumérien, au phénicien, à l’araméen, etc... La plupart de ces langues sont aujourd'hui disparues, à l'exception du copte, resté une langue liturgique, et de l’araméen, parlé par moins de 100 000 locuteurs. De toutes les langues **chamito-sémitiques** actuelles, l’arabe, avec ses variétés dialectales, constitue l’idiome parlé par le plus grand nombre de locuteurs (au moins 200 millions). Avec près de 300 millions de locuteurs dans le monde, les langues de la famille **chamito-sémitique** font partie des familles les plus importantes du monde, tant par leur histoire que par la distribution géographique de leurs locuteurs. L'alphabet arabe moderne et

l'alphabet hébreu se sont développés à partir de la variante araméenne, laquelle a également donné naissance à l'alphabet grec. L'araméen a été supplanté à son tour par l'arabe.

La langue arabe est rattachée à la branche du Sémitique Occidental, sous le groupe Sémitique-Central (voir Figure 4) (Brockelmann, 1910* ; Hetzron, 1976). Ce classement est établi sur la base de recherches morpho-lexicales récentes, au lieu d'étudier les innovations phonologiques, syntaxiques ou lexicales utilisées traditionnellement pour la classification des langues sémitiques.

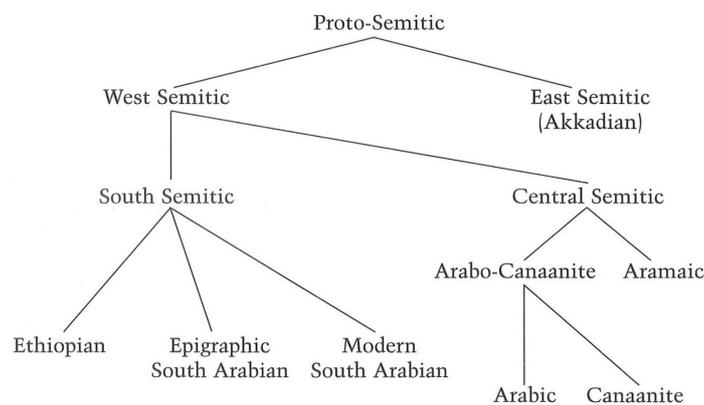


Figure 4 : Classification généalogique des langues Sémitiques telle que établie par Hetzron (1976).
(Emprunté à Versteegh, 2001).

Du point de vue phonologique, une des caractéristiques des langues sémitiques est le nombre important des consonnes par rapport aux voyelles (en arabe classique, on observe la présence de vingt-huit consonnes (voir Tableau 1) contre trois segments vocaliques longs et trois brefs, situés dans les extrémités de l'espace vocalique, voir *infra*).

Pour ce qui concerne le système consonantique, les langues sémitiques présentent une catégorie de consonnes spécifique dite emphatique et qui correspond aux consonnes suivantes $/*t^{\text{e}} *s^{\text{e}} *k^{\text{e}} *k'^{\text{e}} *ð^{\text{e}} *θ^{\text{e}}/$, desquelles la langue arabe n'a conservé que quatre consonnes pharyngalisées ou vélarisées $/t^{\text{e}} d^{\text{e}} s^{\text{e}} ð^{\text{e}}/$ correspondant à $/*t^{\text{e}} *θ^{\text{e}} *s^{\text{e}} *ð^{\text{e}}/$ et une occlusive uvulaire $/q/$ pour $/*k'^{\text{e}}/$ ¹² (Martinet, 1953 ; Bomhard, 1995).

¹² Ces consonnes sont généralement utilisées pour la discrimination des dialectes arabes.

Tableau 1 : Système consonantique de l'arabe classique¹³.
(Adapté de Kaye & Daniel, 1997* ; emprunté à Barkat, 2000)

	Bilabiale	Labiodentale	Interdentale	Alvéolaire	Post-alvéolaire	Pharyngalisé	Palatale	Vélaire	Uvulaire	Pharyngale	Glottale
Occlusive	b			t d		t ^s d ^s		k	q		ʔ
Nasale	m			n							
Trille				r							
Fricative		f	θ ð	s z	ʃ ʒ	s ^s ð ^s		χ	ħ	ʕ	h (ħ)
Approximante							j	w			
Latérale				l							

Cette évolution a été expliquée par différents auteurs. En effet, Roman (1981/82_b) pense que les langues sémitiques, et plus particulièrement l'arabe, ont choisi la constriction pharyngale pour distinguer les consonnes glottalisées sourdes de l'Afro-Asiatique, qui ont perdu leur glottalisation en devenant voisées, constriction produisant une bémolisation du deuxième formant. Ce système a évolué du fait de son organisation propre et du fait de l'organisation des systèmes connexes, syllabique et morphologique. Les consonnes emphatiques de l'arabe sont devenues alors des consonnes pharyngales et pharyngalisées et elles n'ont pas conservé leur caractéristique éjective. Versteegh (2001) défend cette hypothèse en mettant en correspondance les consonnes vélarisées ou pharyngalisées de l'arabe avec les consonnes glottalisées de l'Éthiopien. Il explique en effet que le changement de qualité de ces consonnes de la glottalisation vers la pharyngalisation est plus facile que l'inverse. On considère alors que ce changement en arabe est un changement secondaire (Versteegh, 2001 : p. 20).

Une des questions importantes qui a préoccupé les différents auteurs porte sur la sonorité de qāf arabe. Cantineau (1936*, 1937*, 1938*) explique que la sonorité de qāf est un trait particulier à certains dialectes arabes et ne se retrouve dans aucune langue sémitique dont la phonétique soit connue directement. Pour lui, le qāf arabe est emprunté à l'araméen. Néanmoins, Cohen (1962) et Martinet (1953) expliquent que la koinè arabe poético-coranique

¹³ La consonne /ɣ/ est généralement décrite par les dialectologues soit comme /ɣ/ soit comme /ɣ/ ; la consonne /ħ/ est considéré soit comme une consonne voisée /ħ/ soit non voisée /ħ/.

du 7^{ème} siècle avait pour qāf un correspondant sourd, d'où le /q/ citadin actuel. Autrement, l'évolution du qāf suit le schéma suivant : *k' > *g > k^j > q (Martinet, 1953).

Ghazeli (1977) fait une synthèse de la controverse sur le caractère sourd/sonore du qāf en arabe ancien. Étant donnée la qualité sonore de la production de qāf arabe, différents auteurs (Martinet, 1953, entre autres) ont fait l'hypothèse, soit d'une origine de ces consonnes commune à toutes les langues sémitiques (car cette qualité serait présente dès le proto-sémitique), soit d'une origine couchitique (à cause de la correspondance entre les emphatiques arabes et les éjectives éthiopiennes). Ghazeli s'écarte de l'hypothèse de Martinet sur la préglottalisation de l'éjective (i.e., /*k'/ > /'g/), néanmoins il est d'accord avec lui pour le passage /*k'/ > /q/ (/*k'/ > /*q'/ > /q/). L'auteur pense que la disparition de l'aspiration a un effet sur la perception de la voix, ce qui a pu être interprété comme un voisement par certains locuteurs (d'où le g des bédouins).

Pour ce qui concerne les voyelles, il semblerait que l'arabe a hérité du même système vocalique du Sémitique, à savoir un système vocalique triangulaire constitué de trois phonèmes longs /i: a: u:/, avec leurs correspondants brefs respectifs /i a u/. Différents travaux proposent un système vocalique long du proto-sémitique plus riche avec la présence des deux voyelles /e: o:/. Rabin (1951*) a reconstruit la voyelle /e:/ en proto-sémitique, en s'appuyant sur les données d'une variété de l'arabe ancien de la Péninsule-Arabique : l'arabe occidental. Cowan (1960) expliquent la présence de la voyelle /o:/ en proto-sémitique là où on retrouve un /a:/ en arabe (e.g., *[falo:m] > [sala:m]). Cowan (1960) fait remarqué que le système vocalique du proto-sémitique est plus proche de celui des dialectes modernes de la Péninsule Arabique que de celui de l'arabe classique. Ehret (1995) défend également l'hypothèse de la présence d'un système vocalique à cinq timbres avec une opposition de quantité, en **afro-asiatique**. (Pour des explications plus détaillées, voir Embarki, 2007_b, à paraître).

1.3.2. Structure de la langue arabe

La langue arabe présente une variété de registres linguistiques qui cohabitent quelquefois ensemble. Les registres linguistiques sont représentés par différentes formes

existantes de la langue, telles que : l'arabe **classique**, **littéral** ou **littéraire** (variété haute), par opposition à l'arabe **dialectal** (variété basse), qui coexistent en variétés régionales. Entre ces deux formes existe une variété écrite et parlée désignée par l'**arabe moderne**, **vivant** ou encore l'**arabe standard moderne** ou **contemporain**. Cette coexistence de différents registres linguistiques conduit à des mélanges d'utilisation par les locuteurs arabophones des variétés hautes et basses dans les différentes situations communicationnelles. Différents auteurs ont décrit la cohabitation des différents registres, en parlant de situation **diglossique** (Marçais, 1930-31* ; Ferguson, 1959), **triglossique** (Youssi, 1983*, 1995* ; Ennaji, 1991*, 2002_a) et même **quadriglossiques** (Ennaji & Sadiqi, 1994* ; Ennaji, 2002_a)¹⁴.

Malgré l'existence d'une variété de registres linguistiques importante en arabe et leur utilisation différenciée en fonction des situations communicationnelles, l'arabe **dialectal**, qui est la langue maternelle de tous les locuteurs arabophones, reste avant tout la langue utilisée dans la vie quotidienne locale. C'est une langue acquise dès la petite enfance et non apprise à l'école ou à l'université (par opposition à l'arabe **classique** ou à l'arabe **standard contemporain**).

Différents auteurs ont montré l'existence de différences importantes entre l'arabe **classique** et l'arabe **dialectal**, tant au niveau du lexique, que de la syntaxe, ou de la morphologie (voir par exemple Versteegh, 2001 : pp. 99-100 pour les différences au niveau morphologique). Les différences phonéto-phonologiques sont plus faciles à appréhender et sont d'une grande importance en dialectologie. Nous observons par exemple des divergences dans le traitement différencié des interdentes /θ ð d^h ð^h/, de la consonne uvulaire /q/, de la consonne vélaire /k/, ou dans le traitement différencié des voyelles brèves (qui chutent dans

¹⁴ Ferguson (1959) parle d'une situation **diglossique**, où deux variétés haute et basse, l'arabe classique et l'arabe dialectal respectivement coexistent dans les situations communicationnelles. Youssi (1983*, 1995*) et Ennaji (1991*) expliquent la situation **triglossique** en arabe par la coexistence de trois variétés : l'arabe classique (dans les mosquées, la poésie et littérature classique, etc...), l'arabe standard moderne contemporain (dans les discours politiques ou scientifiques, les médias, l'administration, etc...) et l'arabe dialectal (dans les situations informelles, à la maison, entre amis, etc...). Quant à la situation **quadriglossique**, Ennaji & Sadiqi (1994*) expliquent qu'outre les trois variétés décrites dans la situation triglossique, il existe un registre nommé « l'arabe parlé des intellectuels », qu'ils utilisent dans les situations communicationnelles de tous les jours (Youssi, 1983*, parle de l'arabe médian pour exprimer ce registre).

certaines dialectes et dans certaines positions) et de l'évolution des diphtongues vers des voyelles périphériques longues. (Pour plus de détails sur les différences existantes entre les variétés en arabe, voir les travaux de D. Cohen, 1962 ; Ghazali, 1979 ; Alioua, 1987 ; Abd-el-Jawad & Awwad, 1989 ; Abou-Haidar, 1991 ; Rjaibi, 1993 ; Embarki, 1997 ; Sabhi, 1997 ; Versteegh, 2001 ; etc...).

Bien que l'utilisation de l'arabe **classique** ou de l'arabe **standard contemporain** ait pour but d'unifier progressivement le monde arabe (au moins au niveau de la langue), nous observons des variations importantes entre les dialectes arabes tant au niveau phonéto-phonologique que syntaxique, morphologique ou encore lexical. Différentes études ont proposé d'établir des classifications des dialectes arabes sur la base à la fois des différences géolinguistiques, mais également sociolinguistiques. La limite de la plupart de ces classifications est la prise en compte d'un seul facteur à la fois, i.e., une classification phonéto-phonologique, ou lexicale, etc... et surtout d'un seul dialecte « représentatif » de tous les dialectes dans une zone précise. Dans la partie qui suit, nous proposons une typologie des dialectes arabes.

1.3.3. Typologie des dialectes arabes

La dialectologie arabe propose différentes classifications des dialectes en s'appuyant sur des facteurs géolinguistiques et sociolinguistiques. Dans les parties qui suivent, nous développons ces différentes visions.

1.3.3.1. Division Géographique

Différents atlas dialectaux ont vu le jour pour mettre en évidence les différences dialectales et régionales dans le monde arabophone. Les plus anciens atlas ont été élaborés pour étudier la variabilité régionale des dialectes Syro-Libanais (Bergsträßer, 1915*), des dialectes de la région de Ḥōrān (la zone englobant la Syrie, le Liban, la Palestine et une partie de la Jordanie) et de Palmyre (Cantineau, 1935*), des dialectes Égyptiens (Abul-Fadl, 1961* ; Behnstedt & Woidich, 1985*, 1987*, 1988*), des dialectes du Nord du Yémen (Behnstedt,

1985*, 1992*), ou encore des dialectes de Tunisie (Baccouche & Merji, 2000_a*, 2000_b*, 2000_c*, 2000_d*).

Ces différentes études ont proposé des classifications d'une zone précise englobant une ou plusieurs aires dialectales et régionales. Il existe également d'autres études qui tentent de classifier les dialectes arabes en différentes zones géographiques aboutissant à des résultats hétérogènes. En effet, certains auteurs parlent de cinq zones régionales et dialectales, d'autres de deux ou trois zones, nous y reviendrons un peu plus loin. Toutefois, la plupart des études sont d'accord pour exprimer la division en cinq zones dialectales suivantes, que nous développerons rapidement (Versteegh, 2001) :

- Les dialectes de la Péninsule Arabique,
- Les dialectes Syro-Libanais ou les « parlers levantins » ou de « bilād iššām »,
- Les dialectes Mésopotamiens,
- Les dialectes Égyptiens, et enfin :
- Les dialectes Maghrébins.

ZONE 1. LES DIALECTES DE LA PÉNINSULE ARABIQUE

Les descriptions linguistiques de cette zone ne sont pas très développées (voir cependant les travaux de Cantineau, 1936*, 1937* ; de Rabin, 1951* ; de Johnstone, 1967_a*, 1967_b ; d'Ingham, 1971, 1982* ; de Holes, 1990* ; de Palva, 1991* ; etc...). Rabin (1951*) explique que pendant la période préislamique, une distinction entre les parlers de l'Est et de l'Ouest de la Péninsule pouvait exister, néanmoins, les déplacements très fréquents des bédouins dans cette zone ont aidé à la diversification très importante des caractéristiques de ces parlers. D'après Versteegh (2001), tous les parlers de cette zone sont des parlers de nomades. Ingham (1982*) et Palva (1991*) distinguent dans cette zone quatre groupes dialectaux :

- G1. Les parlers Arabiques du Nord-Est : regroupant les dialectes du Nejd (divisés en trois sous-groupes : i) les parlers ʿanazi (du Koweït, de Bahreïn (parler des sunnites) et des autres pays du Golfe), ii) les parlers Šammar (incluant certains parlers bédouins

d'Irak) et iii) les parlers bédouins Syro-Mésopotamiens (parlers de bédouins du Nord d'Israël et de la Jordanie).

- G2. Les parlers Arabiques du Sud-Ouest : regroupant les parlers du Yémen, de Hadramaout, d'Aden, ainsi que les dialectes chiites de Bahreïn.
- G3. Les dialectes Arabiques de l'Ouest (domaine Ḥiğāz) : regroupant les parlers de nomades du Ḥiğāz (i.e., littoral de la Mer Rouge et du Golfe d'Aqaba) et de Tiḥāmā (i.e., Nord Yémen) et les parlers de sédentaires des grands centres urbains comme la Mecque et Médine.
- G4. Les dialectes Arabiques du Nord-Ouest : regroupant les parlers bédouins du Néguev, du Sinaï, du Sud de la Jordanie et de la côte occidentale du Golfe d'Aqaba et des régions du Nord-Ouest de l'Arabie Saoudite.

Du point de vue du système vocalique de cette zone, Johnstone (1967_a*, 1967_b) parle d'un système vocalique à huit voyelles /i i: e: a a: o: u u:/ (/e: o:/ correspondent aux anciennes diphtongues de l'arabe classique /aj aw/ respectivement).

ZONE 2. LES DIALECTES SYRO-LIBANAIS

Beaucoup de descriptions ont été proposées pour les dialectes de cette zone (voir les travaux de Barthélemy, 1935-69* ; Cantineau, 1938* ; Ferguson, 1956* ; Cleveland, 1963 ; Fleisch, 1974* ; Iriksoussi, 1981* ; Borg, 1984* ; Suleiman, 1985* ; Palva, 1984, 1991* ; Lentin, 1994, 1995/96* ; etc...). La plupart des classifications propose une subdivision des dialectes Syro-Libanais ou des « parlers levantins » ou de « bilād iššām », en trois sous-groupes :

- G1. Les parlers libanais et du Centre syrien : regroupant la majorité des dialectes libanais, les parlers syriens, et le dialecte "maronite" de Chypre, généralement rattaché aux dialectes libanais (Borg, 1984*).
- G2. Les parlers syriens du Nord : comme, par exemple, celui d'Alep.
- G3. Les parlers syriens du Sud : ce groupe englobe les parlers citadins et villageois de la Palestine Centrale, les parlers du Sud de la Palestine et certains parlers jordaniens du sud de la Syrie, dans la région du Ḥōrān dont une classification détaillée a été produite par Palva, 1984, 1991*).

D'une façon générale, le système vocalique dans cette région est un système à huit voyelles /i i: e: a a: o: u u:/. Il existe néanmoins des variations importantes au niveau de la qualité vocalique. À Tripoli par exemple, Fleisch (1974*) parle d'un système à sept voyelles /i i: e: a o: u u:/ où le /a:/ a fusionné avec le /o:/. Dans les parlers des Druz du Sud du Liban, le /a:/ subi un phénomène d'imāla très important et devient alors /i:/ (Cantineau, 1938*). Pour Ferguson (1956*) et Iriksoussi (1981*), le dialecte de Damas peut connaître un système à dix voyelles avec cinq timbres vocaliques opposant les voyelles longues et courtes : /i i: e e: a a: o o: u u:/.

ZONE 3. LES DIALECTES MÉSOPOTAMIENS

Les dialectologues expliquent que la plupart des parlers appartenant à cette zone sont de type bédouin (voir Blanc, 1964* ; Ingham, 1973, 1976 ; Jastrow, 1978*, 1979*, 1981*, 1990*, 1994* ; etc...). Les caractéristiques communes entre ces différents parlers peuvent contribuer à une classification des parlers de cette zone en un seul et unique groupe en Irak et en Turquie.

Au niveau vocalique, les différents parlers de cette zone présentent un système vocalique à cinq voyelles longues /i: e: a: o: u:/ et à deux voyelles brèves /ə a/, où /ə/ est la forme réduite de /i u/. Le phénomène d'Imāla est largement attesté dans cette aire dialectale.

ZONE 4. LES DIALECTES ÉGYPTIENS

Les études dialectologiques de la zone égyptienne sont abondantes (voir Lethem, 1920* ; Hagège, 1973 ; Kaye, 1976*, 1982*, 1986* ; Woidich, 1979*, 1980*, 1990*, 1993* ; Owens, 1993* ; Roth, 1994 ; etc...). Les différentes classifications des parlers égyptiens proposent une division en quatre sous-groupes :

- G1. Les parlers de la région du Delta : se subdivisent eux-mêmes entre parlers de l'Est et parlers de l'Ouest. Ces derniers partagent quelques traits morphologiques avec les parlers Maghrébins, et sont situés par conséquent dans la zone intermédiaire,
- G2. Le parler du Caire,

- G3. Les parlers égyptiens « intermédiaires » : Ils correspondent aux parlers de moyenne Égypte (i.e., « MittelÄgypten » de Woidich, 1980*), et rassemblent tous les parlers situés depuis le plateau de Gizeh (au Sud du Caire) jusqu'à Assiout.
- G4. Les parlers de Haute-Égypte : englobant tous les parlers dans les régions situées au Sud d'Assiout jusqu'à la frontière soudanaise.

Au niveau du système vocalique, tous les parlers égyptiens présentent un système vocalique à cinq voyelles longues /i: e: a: o: u:/ et à trois voyelles brèves /i a u/, mais les deux voyelles /i u/ sont élidées en syllabes ouvertes non accentuées.

ZONE 5. LES DIALECTES MAGHRÉBINS

Les études dialectologiques dans cette zone sont abondantes (voir M. Cohen, 1912* ; D. Cohen, 1963*, 1964*, 1975* ; Grand'Henry, 1972* ; P. Marçais, 1977 ; Singer, 1984* ; Talmoudi, 1980*, 1981*, 1984* ; Owens, 1984* ; Caubet, 1983, 1993, 1998* ; Taine-Cheikh, 1978*, 1991* ; Laria, 1996 ; etc...). La caractéristique commune aux dialectes appartenant à cette zone est que le processus d'arabisation s'est déroulé en deux stades distincts. Ceci a contribué à la complexification des structures linguistiques des parlers de cette zone. En effet, les autochtones de cette zone dialectale ne sont pas des arabophones mais des populations berbérophones. Leur arabisation a commencé lors des premières invasions de l'Afrique du Nord pendant la seconde moitié de 7^{ème} siècle. Les « envahisseurs » arabes se sont alors installés dans les centres urbains déjà existants à cette époque. Versteegh (2001) explique que cette situation a contribué au développement de formes vernaculaires de type « sédentaire », dont les parlers arabes des Juifs du Maghreb constituent de bons exemples (voir également Cohen, 1912*). Durant cette période d'arabisation nommée la période pré-hilalienne, seules les zones urbaines importantes furent arabisées, les zones rurales et bédouines du Maghreb quant à elles restent à grande majorité berbérophone.

Aux 10^{ème} et 11^{ème} siècles, la seconde phase d'arabisation de tout le Maghreb a pris place pendant les invasions perpétrées par les Banū Hilāl. Pendant cette seconde période, la période hilalienne, la langue arabe a été introduite dans les zones rurales et nomades de l'Afrique du Nord.

Les dialectologues montrent que ces deux phases d'arabisation ont contribué à l'élaboration de deux types de parlers : les parlers sédentaires de la période pré-hilalienne (présents dans les grandes villes), et les parlers bédouins de la période hilalienne. On distingue les groupes suivants :

- G1. Les dialectes « pré-hilaliens » de l'Est parlés en Libye, en Tunisie, et à l'Est-algérien, ayant comme caractéristique la conservation des trois voyelles brèves de l'arabe classique /i a u/,
- G2. Les dialectes « pré-hilaliens » de l'Ouest parlés dans l'Ouest Algérien et au Maroc et se caractérisent par la présence de seulement deux voyelles brèves /ə u/,
- G3. Les parlers « hilaliens » des Sulaym de l'Est : avec parlers du Sud Tunisien et du désert libyen,
- G4. Les parlers « hilaliens » de l'Est : regroupant les dialectes du Centre tunisien et de l'Est algérien,
- G5. Les parlers « hilaliens » du Centre : avec les parlers algériens du Centre et ceux du Sud.
- G6. Les parlers « hilaliens » de Ma'qil : regroupant les parlers de l'Ouest du Maroc, de l'Algérie, ainsi que le dialecte arabe de Mauritanie.

Au niveau vocalique, tous les dialectes Maghrébins, hormis les dialectes de l'Est (de la Libye, de la Tunisie, et de l'Est-algérien) se distinguent des autres variétés de l'arabe par la présence de deux voyelles brèves /ə u/, où /ə/ et la conséquence directe de la réduction de /i/ et /a/. En plus des voyelles brèves, tous les dialectes Maghrébins présentent les trois voyelles longues héritées de l'arabe classique /i: a: u:/. Les anciennes diphtongues /aj aw/ ont convergé vers /i: u:/ respectivement.

Versteegh (2001) explique que cette classification en cinq zones dialectales est basée en grande partie sur des facteurs purement géographiques, car la plupart de temps aucune autre explication n'est avancée :

« The usual classification of the Arabic dialects distinguishes the following [five] groups [...] It is not always clear on what criteria this current classification is based. In some cases, purely geographical factors may have influenced the classification (e.g., the Arabian Peninsula). »

(Versteegh, 2001 : p. 145).

C'est pourquoi, d'autres classifications des dialectes arabes en deux voire trois zones dialectales ont été proposées. En effet, différents auteurs distinguent deux grandes zones dialectales : les dialectes du Moyen-Orient (tous les pays arabes de l'Asie, l'Égypte, le Soudan et l'Est Libyen) et les dialectes du Maghreb (le Maroc, l'Algérie, la Tunisie et l'Ouest Libyen), (voir par exemple les études de M. Cohen, 1912* ; D. Cohen, 1962 ; Cantineau, 1936*, 1937*, 1938*, 1960 ; P. Marçais, 1977 ; Versteegh, 2001 ; etc...). Au niveau rythmique, certains travaux proposent de considérer une troisième zone intermédiaire composée de la Tunisie et de l'Égypte (voir par exemple Ghazali *et al.*, 2002 ; etc...). Il reste néanmoins clair que la division géographique en deux zones, i.e., dialectes du Moyen-Orient *vs* dialectes du Maghreb est la plus admise par les dialectologues arabes.

La comparaison entre les deux classifications en deux et cinq zones révèle que le groupe des dialectes maghrébins reste homogène dans les deux cas, tandis que celui des dialectes du Moyen-Orient est composé quant à lui de quatre groupes dialectaux différents. La justification de la littérature quant à la division en deux zones dialectales seulement est le fait qu'au niveau phonético-phonologique, et plus particulièrement au niveau du traitement des systèmes vocaliques des dialectes arabes, nous observons une tendance à la diminution du nombre de voyelles en passant de l'Est à l'Ouest (P. Marçais, 1977) et plus particulièrement à partir de la Libye. La littérature explique que dans certains parlers maghrébins, et plus particulièrement au Maroc, les voyelles brèves deviennent de simples voyelles ultra-brèves (Benkirane, 1998*). De ce fait la structure syllabique des parlers maghrébins a elle-même été modifiée et présente ainsi des caractéristiques particulières (D. Cohen, 1962 ; Benkirane, 1981/2, 1998*, 2000).

Ces classifications géographiques apportent des éclaircissements importants quant à la situation linguistique de la langue arabe. Toutefois, il ne faut pas oublier dans ces classifications un facteur très important : la coexistence des dialectes sédentaires et bédouins dans toutes les aires dialectales arabes. Dans la partie qui suit, nous nous intéressons à la situation sociolinguistique du monde arabe.

1.3.3.2. Division Sociolinguistique

Dans les premières périodes de l'empire islamique, les dialectes bédouins étaient considérés comme les « seuls vrais représentants » de l'arabe classique. Les bédouins se considéraient alors comme les seuls parlant un « arabe pur » et ce dès les premières périodes de l'islam. En effet, dans la terminologie pré-islamique, une distinction a été établie entre ʿarab, soit les arabes sédentaires vivants dans les villes comme à la Mecque ou à Médine, et ʾaʿrāb, soit les arabes bédouins. Il semblerait que dans la période pré-islamique, lorsqu'on parlait de ʾaʿrāb, ceci indiquait les arabes vivants dans la Péninsule Arabique, tandis que l'appellation ʿarab indiquait les populations sédentaires ou nomades vivant à l'extérieur de la Péninsule Arabique.

À l'intérieur de la Péninsule Arabique, on procéda à une division géographique et ethnique. On observe la présence de deux types d'arabe corrélés à cette dichotomie d'arabe « pur » ou non. Les premiers se font appelés les « vrais arabes » : al-ʿarab al-ʿāriba, qui sont les descendant des tribus de qaḥṭān, et sont considérés comme les arabes du Sud. Tandis que les membres de second groupe ont été, semble-t-il, arabisés tardivement et sont appelés al-ʿarab al-mutaʿarriba ou al-ʿarab al-mustaʿriba. Les premiers font parties des tribus bédouines qui conservent une prononciation « pure » de l'arabe, tandis que les seconds présentent des caractéristiques innovatrices de la langue. Pour Ibn Khaldoun (14^{ème} siècle*), les arabes bédouins parlaient l'arabe selon leurs intuitions linguistiques et n'avaient pas besoin de grammairiens arabes pour leur apprendre la « bonne prononciation » de la langue (voir également Versteegh, 2001, Chapitre 4). Cantineau (1938) explique que :

« La dualité sociologique du Monde Arabe a naturellement son reflet dans la langue : il y a des parlers de sédentaires et des parlers de nomades. Une des questions capitales qui se pose pour le dialectologue est d'établir une discrimination entre ces deux types de parlers et de définir les faits de phonétique, de morphologie, de syntaxe et de vocabulaire qui les opposent ».

(Cantineau, 1938* : p. 80).

Néanmoins, malgré la conservation par certaines tribus bédouines des formes de l'arabe classique, les différentes populations bédouines et sédentaires de la région ne parlaient plus l'arabe classique, mais plutôt une forme mâtinée des caractéristiques des deux parlers. Rabin (1955*), entres autres, explique que l'arabe classique était devenu une langue purement littéraire à la fin de califat omeyyade (661-750).

Au départ, les populations arabes parties pour la conquête islamique étaient constituées de bédouins. Au fur et à mesure de l'invasion de nouvelles villes, les populations sédentarisées ont subi des formes d'innovations et ont vu leur langue se modifier pour s'adapter aux nouvelles populations avec lesquelles elles partageaient les territoires. Les arabes bédouins restés sur place n'ont pas vu leur langue se modifier : ils ont conservé beaucoup de structures de la langue classique qui les distinguent des populations sédentarisées. (Pour les différences entre les parlers de bédouins et de sédentaires, voir Ibn Khaldoun, 14^{ème} Siècle* ; Colin, 1920 ; W. Marçais, 1930/31* ; Cantineau, 1938* ; D. Cohen, 1963*, 1973* ; P. Marçais, 1975 ; Taine-Cheikh, 1978*, 1991* ; Ingham, 1973, 1976, 1982* ; Rosenhouse, 1984* ; Cadora, 1992* ; Vanhove, 1995_a*, 1995_b*, 1996_a, 1996_b*, 1997*, 2002 ; Brett, 1995 ; Lévy, 1995* ; Caubet, 1998* ; Versteegh, 2001 ; etc...).

La présence de ces deux types de parlers : bédouins et sédentaires est liée aux vagues d'arabisation du monde arabe. En effet, les études dialectologiques arabes mentionnent toutes la présence de deux vagues d'arabisation : la période pré-hilalienne et la période hilalienne (Marçais, 1961 ; Singer, 1994* ; Kallas, 1999 ; Versteegh, 2001 ; etc...). Pendant la première phase d'arabisation : la période pré-hilalienne, ce sont les populations bédouines de la Péninsule Arabique qui ont contribuées à l'arabisation des différentes tribus présentes dans les autres parties du Monde Arabe, ce qui a conduit à la création des populations sédentarisées dans les grandes villes conquises. Les autres populations bédouines qui n'ont pas participé à cette phase d'arabisation étaient présentes à l'extérieur des centres urbains et ont conservé une forme de vie non sédentaire. La sédentarisation de ces populations nomades et le contact avec les langues des populations locales a affecté les dialectes bédouins et a contribué à une différenciation importante entre les dialectes sédentaires et bédouins.

Pour ce qui concerne la seconde phase d'arabisation : la période hilalienne, ce sont différentes tribus nomades telles que les Banū Hilāl, les Sulaym ou les Ma'qil qui ont contribué à l'arabisation du reste du Monde Arabe, et par conséquent à la création des dialectes ruraux et nomades. D'après certaines théories, les nouvelles populations arabes ont contribué à l'homogénéisation de la situation linguistique de l'empire islamique, car à l'inverse des dialectes urbains, leurs dialectes n'ont pas subi les effets des contacts avec les dialectes et langues des populations locales (Diem, 1978*).

La différence entre les deux phases se situe au niveau des types de parlars résultants et de la durée d'accomplissement de l'arabisation. En effet, durant la première phase d'arabisation, ce sont essentiellement les parlars de sédentaires qui se sont répandus sur toute l'aire géographique arabe par opposition aux parlars ruraux et bédouins pendant la seconde. La durée nécessaire à l'arabisation des différentes villes est différente d'une région à l'autre. Dans les régions syrienne et mésopotamienne, les dialectes sédentaires et urbains coexistaient déjà durant les périodes préislamiques et la grande majorité des dialectes bédouins est parlée par des populations ayant des contacts avec les populations bédouines de la Péninsule Arabique.

Dans tous les pays de l'Afrique du Nord, une distance chronologique d'au moins quatre siècles séparent les deux vagues d'arabisation, ce qui peut expliquer le manque de conservation des caractéristiques initiales de ceux-ci dans les dialectes bédouins de l'Afrique du Nord : ce sont des tribus déjà influencées par les parlars de sédentaires et ayant par conséquent modifié leurs parlars, Versteegh (2001) explique également que :

« The chronological distance may also explain why in spite of the different origin of the sedentary and the Bedouin dialects it is still possible to speak of a dialect area in North Africa and Egypt. All dialects of North Africa, for instance, exhibit the central feature of the North African dialects, the prefix n- of the first person singular of the imperfect verb. These dialects arrived when there were already prestigious cultural and political centres; and although the Bedouin represented the new military power in the region, they could not avoid the centripetal influence of the sedentary dialects ».

(Versteegh, 2001 : pp. 145-146).

Comme nous l'avons explicité, il existe différentes classifications des dialectes arabes, qui se basent sur des critères géolinguistiques et sociolinguistiques. Il est vrai que la première, géolinguistique, est plus facile à appréhender par rapport à la seconde, sociolinguistique : une division délimitée dans la plupart des cas par des frontières entre pays ou région. Au niveau sociolinguistique, ce sont des parlers qui coexistent dans une même région et entre des villes voisines, et donc plus difficiles à délimiter qu'au niveau de chaque pays.

Les populations arabes actuelles (bédouines, sédentaires, rurales) n'ont pas cessé de modifier leurs habitats au cours des siècles. La plupart des bédouins se sont sédentarisés (il reste néanmoins quelques tribus bédouines non sédentarisées dans la Péninsule Arabique). En dialectologie arabe, on est en présence de parlers citadins (les parlers sédentaires des villes) et de parlers de bédouins (des populations nomades ou transhumantes) auxquels s'ajoutent les parlers villageois (les parlers ruraux et montagnards) (Marçais, 1975).

Les études sociolinguistiques actuelles sur les mouvements de populations arabes et les mélanges de parlers affirment la présence d'une trichotomie, au moins au Maroc : parlers citadins, parlers urbains et parlers ruraux (Messaoudi, 2002). L'auteur explique que les **parlers ruraux** sont utilisés essentiellement par des populations vivant en dehors des villes (dans les plaines, les montagnes), les **parlers urbains** par des populations vivant dans les espaces urbains et enfin les **parlers citadins** par des populations autochtones (en Maghreb se sont essentiellement les parlers aux caractéristiques pré-hilaliennes). Messaoudi (2002) explique que les parlers **citadins** sont en nette régression, les parlers **ruraux** sont supposés vivants et confinés à un espace géographique délimité, les parlers urbains sont quant à eux en **développement** (Messaoudi, 2002 : pp. 225).

Une autre vision intéressante de la complexité sociolinguistique du Monde Arabophone est proposée par Embarki (2007_a, 2007_b). Pour l'auteur, quinze systèmes linguistiques différents sont présents dans le monde arabophone : cinq zones géographiques et trois situations sociolinguistiques (sédentaires citadins, sédentaires villageois et nomades), avec des caractéristiques précises à chaque système (voir Tableau 2).

À l'inverse des autres dialectologues arabes pour qui les parlers de la ville (i.e., parlers citadins) sont mobiles et en perpétuel développement tandis que les parlers de la campagne sont fixes, Messaoudi (2002) et Embarki (2007_a, 2007_b) proposent trois situations sociolinguistiques différentes : les **parlers citadins** et **ruraux** sont caractérisés par la fixité (i.e., en régression, sans évolution, confiné dans des zones précises, etc...), tandis que les **parlers urbains** sont mobiles et en nette progression.

Tableau 2 : Liste des unités phonologiques, vocaliques et consonantiques, caractéristiques des quinze systèmes dialectaux, en fonction des cinq groupes régionaux et des variétés sociologiques qui les divisent. (Emprunté à Embarki, 2007_b).

Division géographique/ sociologique	bédouins nomades	bédouins sédentaires	citadins
Arabique	dz-dʒ, θ, ð, ð ^ʕ , i:, u:, e:, o:, a:, i, u, a	dʒ-g, θ, ð, ð ^ʕ , i:, u:, e:, o:, a:, i, u, a	dʒ-g, θ, ð, ð ^ʕ , i:, u:, e:, o:, a:, i, u, a
Mésopotamien	dʒ, θ, ð, ð ^ʕ , i:, u:, e:, o:, a:, i, u, a	g, θ, ð, ð ^ʕ , i:, u:, e:, o:, a:, i, u, a	q, t, d, d ^ʕ , i:, u:, e:, o:, a:, i, u, a
Levantin	k, θ, ð, ð ^ʕ , i:, u:, e:, o:, a:, i, u, a	k ^ʕ -g, θ, ð, ð ^ʕ , i:, u:, e:, o:, a:, i, u, a	ʔ, t-s, d-z, d ^ʕ -z ^ʕ , i:, u:, e:, o:, a:, i, u, a
Égyptien	g, s, z, z ^ʕ , i:, u:, e:, o:, a:, i, u, e, o, a	k, s, z, z ^ʕ , i:, u:, e:, o:, a:, i, u, e, o, a	ʔ, s, z, z ^ʕ , i:, u:, e:, o:, a:, i, u, e, o, a
Maghrébin	g, θ, ð, ð ^ʕ , i:, u:, a:, i, u, a, ə	g, t, d, d ^ʕ , i:, u:, a:, i, u, a, ə	q, t, d, d ^ʕ , i:, u:, a:, i, u, a, ə

La situation linguistique de la langue arabe que nous venons de présenter est donc très complexe. Nous avons vu que souvent la littérature n'est pas unanime pour telle ou telle classification. Toutefois, nous observons dans la dialectologie arabe actuelle une forte orientation vers les études portant sur la classification sociolinguistique : Suleiman (1996, 2003) explique en effet que **pour un arabophone, ce qui compte, ce n'est pas ce que nous étions, mais ce que nous voulons être (notre identité)**. Nous cherchons avant tout à adopter le système de ceux à qui nous voulons ressembler. C'est pourquoi, nous observons de plus en plus de modifications des parlers de jeunes populations arabes : ceux-ci modifient leurs parlers en empruntant l'« accent » d'un parler plus prestigieux ; i.e., les **parlers urbains** qui sont en nette progression. Les divisions géographiques sont importantes mais au-delà de celles-ci, les divisions sociolinguistiques l'emportent.

Dans les parties précédentes, nous avons développé la situation linguistique de l'arabe en abordant son histoire, sa structuration et son organisation. Nous avons ensuite proposé une typologie des dialectes modernes au niveau géolinguistique et sociolinguistique. Nous avons observé les grandes différences existantes entre les différents dialectes en fonction des zones géographiques. Dans cette thèse, nous avons décidé de limiter la comparaison trans-dialectale à deux dialectes qui se situent dans les deux extrémités du Monde Arabe : l'arabe jordanien et l'arabe marocain, dont nous développons les caractéristiques dans les parties qui suivent.

1.4. DESCRIPTION DES DEUX DIALECTES ARABES ÉTUDIÉS

Dans cette partie, nous présentons un aperçu des dialectes arabes étudiés dans cette thèse : la variété dialectale jordanienne, et plus particulièrement le parler d'Irbid (une ville du Nord), qui fait partie du domaine moyen-oriental, et ensuite le dialecte marocain représenté par deux parlers : les parlers de Casablanca et de Fès¹⁵. Nous présentons pour chaque dialecte la situation linguistique et le système vocalique.

1.4.1. Arabe Jordanien

1.4.1.1. Généralités

Du point de vue géographique, la Jordanie est située entre la Syrie au Nord, l'Irak au Nord-Est, l'Arabie Saoudite au Sud et l'état d'Israël à l'Ouest. Elle fait partie de deux régions géographiquement et linguistiquement distinctes :

- À l'Ouest, on trouve la partie de la Cisjordanie, qui comprend tout le Nord de la Jordanie et la partie orientale de la dépression de Ghor jusqu'au Golfe d'Aqaba à la frontière de l'Égypte et d'Israël,

¹⁵ Deux parlers de l'arabe marocain ont été considérés car les locuteurs de Casablanca ont participé à la phase de production des voyelles, tandis que les locuteurs de Fès ont participé à la phase de perception des voyelles en **statique** et en **dynamique**.

- À l'Est, on trouve le plateau de Transjordanie, qui couvre les trois quarts du territoire jordanien et fait suite au désert de Syrie.

La population jordanienne, qui compte à peu près 5.760.000 habitants (estimation 2005), est composée de différentes populations d'Arabes (musulmans sunnites) citadins, bédouins, bédouins semi-sédentarisés et d'autres minorités turques, circassiennes, kurdes et druzes. Ces derniers sont dispersés sur tout le territoire jordanien et contribuent à l'introduction de nouvelles variétés régionales.

L'arabe jordanien est une variété dialectale « basse » (suivant les termes diglossiques, Ferguson, 1959) qui n'est ni codifiée ni standardisée. La grande majorité des travaux en dialectologie jordanienne s'est concentrée sur l'étude des dialectes de bédouins et par conséquent a considéré les parlers jordaniens comme des parlers de bédouins (Czapkiewicz, 1960* ; Cleveland, 1963 ; Palva, 1969_a*, 1969_b*, 1976*, 1978*, 1980*, 1984, 1986*, 1989*, 1993* ; Bani-Yassin & Owens, 1984, 1987* ; Irshied, 1984* ; Irshied & Kenstowicz, 1984* ; Sakarna, 1997*, 1999, 2005 ; etc...). Néanmoins, les populations palestiniennes et les différentes minorités ont contribué à l'évolution du parler jordanien. En effet, selon Palva (1984), les parlers de **bédouins**, de **sédentaires citadins** et de **villageois** coexistent en Jordanie (voir également Abd-el-Jawad, 1981*, 1986* ; Al-Khatib, 1988 ; Abd-el-Jawad & Awwad, 1989* ; Hussein & El-Ali, 1989* ; Al-Sughayer, 1990 ; pour qui ce sont des parlers de **bédouins**, d'**urbains** et de **ruraux**, respectivement). Généralement, ces différents auteurs utilisent le terme **parler urbain** pour désigner le parler des jordaniens d'origine palestinienne qui est différent, selon eux, des **parlers ruraux** des jordaniens d'origine palestinienne et des **parlers de bédouins** et de **ruraux** jordaniens (voir Abd-el-Jawad, 1981*, 1986* ; Abd-el-Jawad & Awwad, 1989 ; Hussein & El-Ali, 1989* ; pour les critiques de ces classifications, voir Sakarna, 2005 ; etc...).

Ces différents parlers forment un continuum linguistique assez particulier qui s'ajoute à la densité linguistique liée aux différentes populations y vivant. La dialectologie arabe décrit la Jordanie comme appartenant à deux zones dialectales distinctes à l'intérieur de la zone Moyen-Orientale : la zone Syro-Libanaise contenant une majorité de parlers de sédentaires

citadins et de villageois et la zone de la Péninsule Arabique contenant une majorité de parlers de bédouins (voir Figure 5).

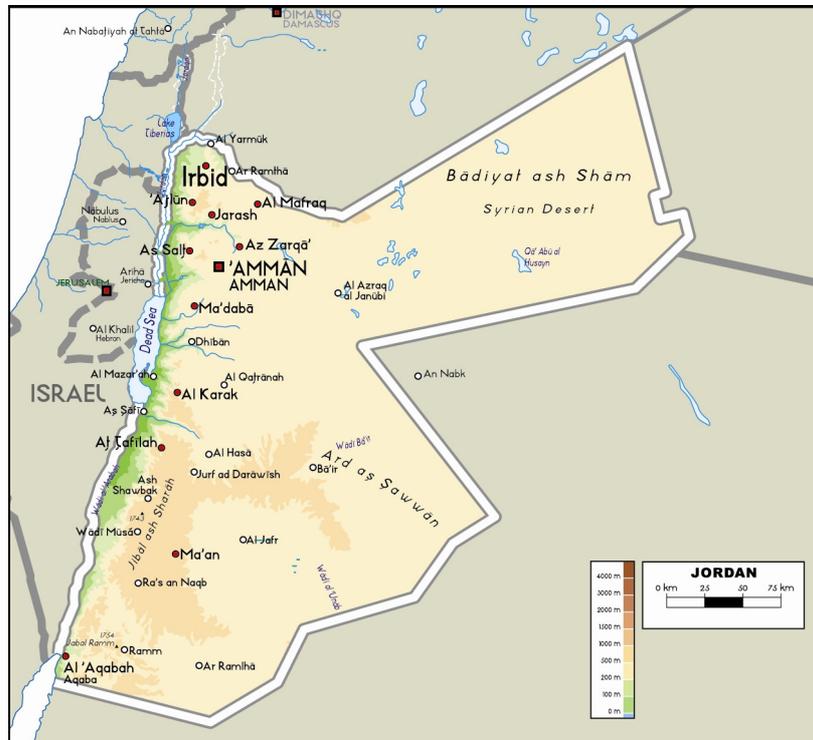


Figure 5 : Carte géographique de la Jordanie, la zone dialectale Syro-Libanaise se situe du nord de la Jordanie jusqu'à la ville d'Al-Karak, la zone dialectale de la Péninsule Arabique se situe dans le désert syrien et au sud. (Adapté de GEOATLAS, 2002).

La différenciation des dialectes jordaniens se fait principalement par le biais des variations consonantiques. En effet, les parlers de bédouins et de villageois conservent la réalisation des fricatives interdentes /θ ð ðʕ/ qui deviennent [t d dʕ] ou [s z zʕ] dans les parlers de citadin ainsi que la réalisation sonore du /q/ qui devient [g] chez les bédouins, [k ou kʰ] chez les villageois et devient [ʔ] chez les citadins, (d'après Cleveland, 1963 ; Palva, 1984 ; Versteegh, 2001 ; etc...).

Différents auteurs ont proposé des classifications des dialectes parlés en Jordanie et en Cisjordanie. Cleveland (1963) a établi une classification en se basant sur un seul critère : les variations morpho-phonologiques de qāf. En effet, il a pu diviser cette région en quatre zones dialectales sur la base de leurs réalisations du mot [yaqu:l] « il dit » :

- Groupe I, ou les dialectes [yigu:l], qui comprend les parlers de bédouin dans l'Est et le Sud de la Jordanie,

- Groupe II, ou les dialectes [bəgu:l], qui comprend les parlers de la région rurale du Sud de la Palestine, la vallée du Jourdain et la plupart des nomades qui n'appartiennent pas au premier groupe,
- Groupe III, ou les dialectes [bəku:l], qui comprend les parlers de villageois autour de Jérusalem et de la Palestine Centrale,
- Groupe IV, ou les dialectes [bəʔu:l], qui comprend les parlers des urbains de Jérusalem, de Jénine et d'Hébron.

Palva (1984), quant à lui a proposé une classification des dialectes parlés dans cette région en utilisant onze critères, dont la réalisation différenciée de qāf et des interdentes¹⁶. Sa classification correspond globalement à celle proposée par Cleveland (1963), mais avec plus de précision, distinguant trois grandes zones dialectales avec des subdivisions (en fonction de la réalisation des items suivants : [yaqu:l] « il dit », [di:k] vs. [dyu:k] « coq » singulier vs. pluriel, [θaqi:l] « lourd » et [qahwah] « café ») :

- La zone des dialectes urbains de la Palestine, avec les formes ⇒ [bəʔu:l],
- La zone des dialectes ruraux :
 - Les dialectes de Galilée, qui sont les dialectes sédentaires les plus anciens, avec les formes ⇒ [biqu:l],
 - Les dialectes de la Palestine Centrale, qui sont les dialectes les plus conservateurs et les plus influencés par les dialectes des bédouins Syro-Mésopotamiens, avec les formes ⇒ [bik^hu:l],

¹⁶ Au total, Palva (1984) a utilisé onze critères de discrimination des dialectes parlés dans la région jordano-palestinienne. Ces critères sont les réalisations différenciées : 1) de qāf (/q/ > [q g ʔ k ou k']), 2) des interdentes (/θ ð δ s/ > [t d d^s] ou [s z z^s]), 3) de kāf (/k/ > [k tʃ ou ttʃ]), 4) de ġīm (/ʒ/ > [ʒ ou dʒ]), 5) de la séquence **CVCaCV-** (CVCaCV- > CVCaCV- ou CCaCV- (ex. /kataba/ « il a écrit » + /aw/ « troisième personne de pluriel » > [katabu] (chez les sédentaires) ou [ktəbaw] (chez les bédouins)), 6) de la séquence -aXC- (-aXC- reste -aXC chez les sédentaires et devient -XaC ou -aXaC chez les bédouins, où X = /x ɣ h ʔ h/, ex. /qahwa/ "café" > [qahwa] chez les sédentaires (voir variation de qāf) et devient soit ['ghawa^h] soit ['gahawa^h] chez les bédouins), 7) de la distinction de deuxième et troisième personne du pluriel, 8) de l'indicatif de l'inaccompli ya- (/y-/ > [b-] chez les sédentaires et [y-] chez les bédouins), 9) de l'adverbe hona "ici", 10) de l'adverbe al-'ān "maintenant", et 11) de la présence ou l'absence du morphème de la négation *maa-*.

- Les dialectes de la Palestine du Sud rattachés aux dialectes de la Palestine du Centre, avec une influence importante des parlers de bédouins, avec les formes ⇒ [bigu:l], [di:t̪] et [dyu:t̪],
- Les dialectes du Nord et du Centre de la Jordanie, qui sont très liés aux dialectes du Ḥōrān. Ils ressemblent beaucoup aux parlers ruraux de la Palestine ainsi que les parlers des bédouins Syro-Mésopotamiens. Ce sont des bédouins sédentarisés, avec les formes ⇒ [bigu:l], [di:t̪] et [dyu:k],
- Les dialectes du Sud de la Jordanie, qui représentent une mixture des dialectes, mais l'influence des parlers bédouins de la région de Pétra est la plus remarquée, avec les formes ⇒ [bigu:l], [di:k], [dyu:k] et [gahwa],

➤ La zone des dialectes bédouins :

- Les dialectes de Néguev, qui sont très proches des dialectes sédentaires du Sud de la Palestine. Ils présentent des similarités avec les dialectes bédouins de Sinaï et de Pétra, avec les formes ⇒ [bigu:l], [di:k], [dyu:k] et [gahawah],
- Les dialectes de Pétra, qui sont des dialectes bédouins aux caractéristiques spécifiques, avec les formes ⇒ [yigu:l], [θagi:l],
- Les dialectes Syro-Mésopotamiens (de la zone des éleveurs de moutons), qui sont des dialectes bédouins proches de ceux de la Syrie, de l'Irak et du Golfe, avec les formes ⇒ [yigu:l], [θid̪ɟi:l],
- Les dialectes des bédouins du Nord (de la zone des éleveurs de chameaux), qui sont des dialectes de bédouins parlés en Jordanie par les tribus Sirḥān, Bani Ṣaḥar et Bani Ḥālīd, avec les formes ⇒ [yigu:l], [θid̪ɟi:l].

La grande majorité des études dialectologiques de la Jordanie décrivent les différences au niveau consonantique et morphologique. Les différents auteurs travaillant dans le domaine jordanien décrivent le système vocalique des différents parlers comme ayant un système à cinq timbres vocaliques, huit voyelles (voir Czapkiewicz, 1960* ; Cleveland, 1963 ; Irshied, 1984* ; Irshied & Kenstowicz, 1984* ; Awwad, 1986* ; Bani-Yassin & Owens, 1987 ; Sakarna, 1997*, 1999, 2005 ; Barkat, 2000 ; Al-Tamimi, 2001, 2002 ; etc...). Ce système est composé des voyelles suivantes : /i i: e: a a: o: u u:/ (les deux voyelles intermédiaires /e: o:/ représentent les anciennes « diphtongues » de l'arabe classique /aj aw/ respectivement) (voir

Figure 6). Néanmoins, il existe des variations importantes au niveau des systèmes vocaliques des dialectes jordaniens :

- Deux voyelles brèves supplémentaires existent dans les parlers urbains (des jordaniens d'origine palestinienne : voir Abd-el-Jawad, 1981*, 1986* ; Abd-el-Jawad & Awwad, 1989 ; et dans le parler bédouin de la tribu des 'ağārma : voir Palva, 1976*),
- Les diphtongues /aj aw/ de l'arabe classique sont toujours réalisées dans les dialectes bédouins des tribus des Bdūl (Bani-Yassin & Owens, 1987 ; et des Bani Şaḥar (Palva, 1980*),
- Réalisations différenciées des voyelles dans la capitale Amman, observées au niveau des populations jeunes et féminines : /ɑ:/ ⇒ [a:] ; /a:/ ou /æ:/ ⇒ [ɛ:] ; /ɛ:/, /e:/ ou /i:/ ⇒ [ɛə ɪə], etc... (voir Al-Wer, 2000, 2002_a, 2002_b).

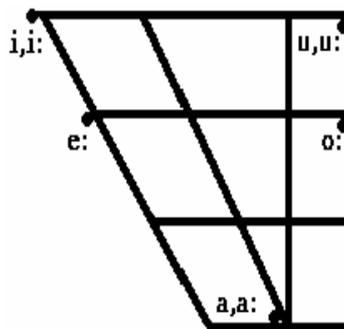


Figure 6 : Système vocalique phonologique de l'arabe jordanien, à cinq timbres, huit voyelles.
(Adapté de Bani-Yassin & Owens, 1987).

Nous nous concentrerons sur l'étude du parler d'Irbid au Nord de la Jordanie qui est considéré comme un dialecte mélangeant les variétés de sédentaires villageois et citadins avec quelques influences des parlers de bédouins.

1.4.1.2. Le parler d'Irbid

Géographiquement, la ville d'Irbid est située au Nord-Ouest de la Jordanie, à l'Est du Jourdain et à proximité de la frontière syrienne. Cette ville, la deuxième ville après la capitale Amman, compte près de 650.000 habitants qui parlent un dialecte mélangeant les caractéristiques des parlers de sédentaires villageois et de citadins. Pour Al-Khatib (1988), deux types de parlers ruraux coexistent à Irbid et dans le Nord de la Jordanie : les Ḥōrānī et

les fallāhīnī ; les premiers sont originaires de la plaine de Ḥōrān dans le Sud de la Syrie, tandis que les seconds sont issus des zones rurales en Palestine et Israël (pour plus de détails sur les différences entre les deux parlers, voir Al-Khatib, 1988 : pp. 349-355). L'auteur explique que :

« the language spoken by the two rural groups in Irbid city is highly varied and that this variation is systematic and rule-governed ».

(Al-Khatib, 1988 : p. 350).

Dans certaines régions du Nord et de l'Est d'Irbid, on observe des influences des caractéristiques des parlers bédouins sur la réalisation des interdentes qui sont réalisées en tant que telles, de la réalisation sonore du qāf /q/ > [g] (sauf pour les termes religieux où la réalisation sourde est conservée) et de l'opposition /k/ ~ [tʃ]¹⁷ qui existe essentiellement dans le Nord de la ville d'Irbid (à proximité de la frontière syrienne) et dans certains parlers ruraux de la région. Ce parler fait partie de la zone des dialectes ruraux de la région jordano-palestinienne, selon la classification de Palva (1984)

Du point de vue du système vocalique du parler d'Irbid¹⁸, Bani-Yassin & Owens (1987) observent la présence de huit voyelles phonologiques /i i: e: a: a: o: u u:/, avec des variations phonétiques en présence de contexte consonantique pharyngalisé (/i i: a: a: u u:/ ⇒ [e e: a: a: o o:]) et de structure syllabique spécifique (de type CVC, /i i: a: a: u u:/ ⇒ [ɪ ɪ: æ æ: ʊ ʊ:]). Les auteurs expliquent la présence de différentes règles d'épenthèse liées à la disparition des voyelles fermées en syllabes ouvertes. D'après les auteurs, ce parler présente un phénomène d'harmonie vocalique lors de la réalisation des voyelles fermées brèves (ex. /yakubbu/ > [bikubb] > [bukubb] « il a jeté », le morphème /ya/ devient [bi] par règle de mutation spécifique aux dialectes moyen-orientaux de la zone Syro-Libanaise, et ensuite, le morphème [bi] subit l'harmonie vocalique et devient [bu]¹⁹). Les différents auteurs travaillant

¹⁷ Le passage de /k/ à /tʃ/ est irrégulier et restreint à certains parlers dans la région d'Irbid. Ce changement est globalement réalisé à l'initial du mot et rarement en position médiane et/ou finale (ex. /ki:s/ > [tʃi:s] "sachet", /kalb/ > [tʃalb] "chien").

¹⁸ Nous ne décrivons pas le système consonantique de ce parler car nous nous intéressons dans ce travail au seul système vocalique. L'arabe jordanien et plus particulièrement le dialecte d'Irbid présente le système consonantique suivant : /b t d k q ʔ m n r f θ ð s z ʃ d̤ ʒ ɣ ħ ʕ h w j l t̤ s̤ d̤ s̤ ð̤/.

¹⁹ Ce phénomène est observable dans certains parlers de villageois de la région d'Irbid.

sur le parler d'Irbid ont remarqué la présence de la voyelle centrale [ə] qui n'est pas conditionnée par un quelconque contexte consonantique ou une structure syllabique particulière (Bani-Yassin & Owens, 1987 ; Barkat, 2000 ; Al-Tamimi, 2001).

1.4.2. *Arabe Marocain*

1.4.2.1. Généralités

Géographiquement, le Maroc est situé dans la zone la plus à l'Ouest de l'Afrique du Nord. Il est délimité à l'Ouest par l'Océan Atlantique, au Nord par le détroit de Gibraltar et la Méditerranée, à l'Est et au Sud par la Mauritanie et l'Algérie (voir Figure 7).

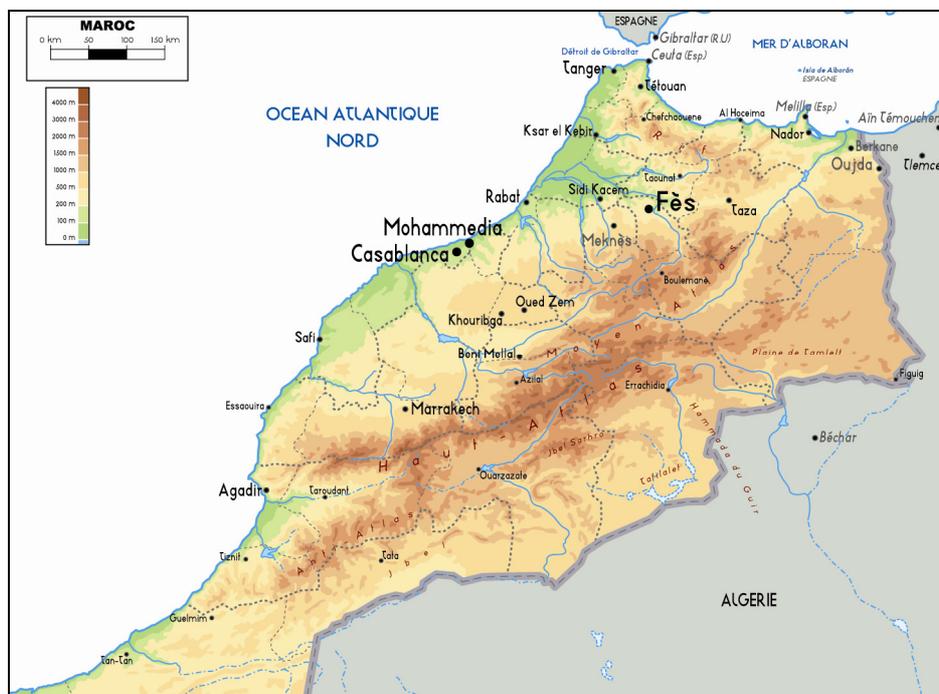


Figure 7 : Carte géographique du Maroc.
(Adapté de GEOATLAS, 2002).

Le Maroc compte actuellement près de 33.241.000 habitants (estimation 2006). Cette population est en contact direct avec la langue arabe (sous sa forme dialectale avant la forme classique), la langue berbère, l'anglais, l'espagnol et le français. Cette situation favorise la complexité linguistique présente sur le territoire : la plupart des marocains parlent ou, au moins, connaissent deux ou trois langues. D'après Ennaji (2002_b), le français prédomine au Maroc surtout dans les domaines de l'éducation, de l'administration, de l'industrie, des

finances et du commerce. Pour lui, l'arabe standard contemporain est utilisé avant tout dans les discours traditionnels et religieux, mais également dans tout ce qui touche à la littérature et aux aspects culturels. Différents auteurs expliquent que la situation du bilinguisme français~arabe standard contemporain observée au Maroc est considérée comme la meilleure option au développement du pays (voir Ennaji, 1988* ; Elbiad, 1995 ; etc...).

L'arabe marocain est une variété dialectale « basse » (suivant les termes diglossiques, Ferguson (1959)) qui n'est ni codifiée ni standardisée. C'est un dialecte parlé par la grande majorité des marocains, arabe et berbère (qui l'apprennent comme une seconde langue (Ennaji, 2002_b)). Concernant les populations arabophones, comme nous l'avons vu, elles ont été arabisées en deux phases : la période pré-hilalienne au 7^{ème} siècle et la période hilalienne au 10^{ème} et 11^{ème} siècle. Comme nous l'avons expliqué, la différence entre les deux phases influence le type de parler résultant. Pour la première, ce sont des parlers sédentaires citadins, tandis que la seconde, des parlers de bédouins. Ces deux phases d'arabisation et les migrations des zones rurales vers les grandes villes ont conduit à des mélanges de parlers au Maroc (comme dans tous les autres pays arabes). Nous observons alors une situation sociolinguistique complexe où parlers de citadins, d'urbains et de ruraux coexistent (voir Messaoudi, 2002 ; Embarki, 2007_a ; 2007_b). Tous les parlers marocains ont subi des influences de sédentaires citadins et villageois et/ou bédouins dues aux migrations rurales vers les grandes villes. Ainsi, on observe une homogénéité de la structure phonologique des parlers arabes marocains et une dominance des formes linguistiques des parlers de grandes villes comme à Casablanca et/ou à Rabat. Outre la variabilité dialectale sociolinguistique, il existe au Maroc une variabilité régionale importante. Différentes classifications opposent les dialectes orientaux (à Tanger, à Tétouan, etc...) aux dialectes occidentaux ou de Ġarb (à Casablanca, à Kénitra, etc...). La plupart des dialectologues distinguent au moins cinq zones dialectales (Ennaji, 2002_b) :

- La zone des dialectes du Nord : les dialectes Šamālī, qui sont parlés à Tanger, à Tétouan, à Larache,
- La zone des dialectes fāsī, qui sont parlés à Fès, à Sefrou, à Meknès,
- La zone des dialectes du Centre, qui sont parlés à Rabat et à Casablanca,

- La zone des dialectes du Sud, qui sont parlés à Marrakech et à Agadir, et enfin
- La zone des dialectes du Sahara, comme le dialecte de Hassaniyya.

En ce qui concerne le système vocalique de l'arabe marocain (voir Figure 8), la littérature est d'accord pour dire qu'il présente trois voyelles « longues » ou « pleines » /i: a: u:/ héritées de l'arabe classique. Les deux diphtongues de l'arabe classique /aj aw/ sont devenues /i: u:/ respectivement. Pour ce qui concerne les voyelles brèves, la littérature n'est pas homogène et propose deux visions différentes.

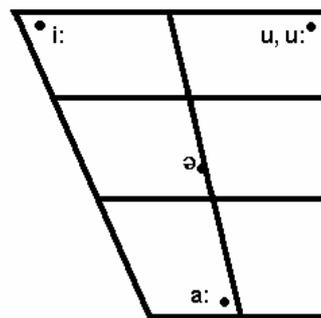


Figure 8 : Système vocalique phonologique de l'arabe marocain, à quatre timbres, cinq voyelles.
(Adapté de Hamdi, 1991).

Les partisans de la première vision considèrent que les anciennes voyelles de l'arabe classique /i a/ ont disparu du système marocain et ont ainsi fusionné pour faire ressortir une seule qualité vocalique brève /ə/. Quant à l'ancienne voyelle /u/, elle conserve son opposition avec la longue /u:/ et avec le /ə/ et devient ainsi /ʊ/. (voir Cantineau, 1960 ; Cohen, 1962, 1965 ; Hamdi, 1991 ; Barkat, 2000 ; etc...).

Pour les seconds, l'opposition de durée vocalique n'existe plus en arabe marocain, les anciennes voyelles longues de l'arabe ont perdu leur qualité de longueur vocalique. Pour ce qui est des anciennes voyelles brèves, vue la perte d'opposition de quantité vocalique, toutes les anciennes voyelles brèves /i a u/ ont fusionnées pour ne former qu'une seule voyelle centrale /ə/ (voir Khomsi, 1975 ; Laabi-Zemmama, 1975* ; Angoujard, 1990 ; Rhardisse, 1995 ; Benkirane, 1982, 2000 ; Dell & Elmedlaoui, 2002 ; etc...). Toutefois, les différents auteurs ne sont pas d'accord pour le statut de cette voyelle centrale : certains supposent qu'elle a un statut phonologique comme les trois autres voyelles pleines (voir Khomsi, 1975 ; Rhardisse, 1995 ; etc...) et d'autres pensent qu'elle n'a qu'une représentation phonétique de

surface uniquement (voir Benkirane, 1982, 2000 ; Dell & Elmedlaoui, 2002 ; etc...). Benkirane (1982, 2000) explique en effet que la voyelle centrale /ə/ n'existe qu'en syllabe fermée CVC, les voyelles pleines sont réalisées en syllabe ouverte, qui prédomine dans le dialecte marocain. L'auteur a observé la présence de cette voyelle centrale pour rompre les suites de plus de trois consonnes.

Dans une étude précédente, (voir Al-Tamimi & Barkat-Defradas, 2002 ; Barkat-Defradas *et al.*, 2003), nous avons montré qu'en production et en perception de la parole, les deux voyelles « brèves » /ə u/ de l'arabe marocain de Casablanca ont été produites et perçues comme une seule et unique voyelle centrale /ə/, uniquement en contexte dental /d d^h/.

Nous pouvons alors nous poser la question de savoir quel est le statut phonologique de cette (ces) voyelle(s) centrale(s) en arabe marocain ? Est-ce pour rompre les suites consonantiques de type CCCV et donc elle(s) n'a (ont) qu'un statut phonétique uniquement ? Il serait alors intéressant d'étudier les incidences de l'évolution du système vocalique de l'arabe marocain du point de vue perceptif, afin de mieux comprendre le statut de ses voyelles.

Dans la partie qui suit, nous nous intéressons aux caractéristiques des deux parlers marocains étudiés dans notre thèse, à savoir l'arabe marocain de Casablanca et celui de Fès.

1.4.2.2. Le parler de Casablanca

Casablanca est la plus grande ville du Maroc. Elle a été fondée par les portugais en 1515 et lui donnèrent le nom de **Casa Branca**, qui a été ensuite hispanisé en **Casa Blanca**. C'est une ville de peuplement tardif, vers les 18^{ème} et 19^{ème} siècles. Elle est la capitale économique du Maroc et son principal port. La ville est située sur la côte Atlantique à environ 90 km au sud de la capitale du pays, Rabat. Elle appartient à la région du Grand Casablanca qui regroupe aussi la ville de Mohammedia (120.000 habitants), Ain Harrouda et les deux provinces Mediouna et Nouasser. La population totale de la région est de 5.600.000 habitants (estimation 2006).

En suivant la classification proposée par Ennaji (2002_b), le parler de Casablanca fait partie de la zone des dialectes du Centre. Différentes études dialectologiques du parler de la ville de Casablanca ont vu le jour (voir par exemple Harris, 1942 ; Cantineau, 1960 ; Cohen, 1962, 1965 ; Khomsi, 1975 ; Caubet, 1993 ; Aguadé & Benyahia, 1990 ; Bouziri *et al.*, 1991 ; Hamdi, 1991 ; Barkat, 2000 ; Al-Tamimi & Barkat-Defradas, 2002 ; Ennaji, 2002_b ; Barkat-Defradas *et al.*, 2003 ; etc...). Caubet (1993) considère le parler de la ville de Casablanca comme un parler de bédouin, néanmoins, la grande majorité des travaux cités sur ce parler le décrivent comme un parler de type sédentaire citadin, à cause des différentes influences sociolinguistiques dues aux migrations successives des autres villes.

Les principales caractéristiques au niveau consonantique²⁰ de ce parler est la conservation de la prononciation sourde du **qaaf** malgré la présence signalée par Cantineau (1960) de la réalisation sonore de cette consonne (i.e., /q/ et [g]). Les anciennes fricatives interdentes de l'arabe classique sont devenues des occlusives alvéolaires : /θ ð ð^ʕ/ ⇒ /t d d^ʕ/. Cantineau (1960) explique également la présence de la paire /r r^ʕ/ ainsi que /z z^ʕ/.²¹

Du point de vue vocalique, les mêmes propositions formulées pour l'arabe marocain s'appliquent à ce parler, à savoir un système vocalique à trois timbres vocaliques /i a u/ plus une ou deux voyelles centrales /ə/ ou /ə u/, respectivement.

1.4.2.3. Le parler de Fès

Fès (ou Fez) est la troisième ville du Maroc après Casablanca et Rabat et compte actuellement près de 1.400.000 habitants. Elle fait partie des quatre villes impériales au Maroc (avec Marrakech, Meknès et Rabat) et est souvent considérée comme le centre spirituel du Maroc. Cette ville a été construite autour des 8^{ème} et 9^{ème} siècles durant la première phase d'arabisation du Maghreb. Cette ville a conservé les caractéristiques de son parler pendant de

²⁰ Comme pour l'arabe jordanien, nous ne décrivons pas le système consonantique car nous nous intéressons dans ce travail au seul système vocalique. L'arabe marocain de Casablanca présente un système consonantique différent de celui de l'arabe classique : /b t d k q ʔ m n r f s z ʒ ʒ̣ ɣ ɣ̣ ɧ ɧ̣ h w j l t^ʕ d^ʕ s^ʕ z^ʕ r^ʕ/.

²¹ La consonne pharyngalisée /z^ʕ/ peut être la résultante de la transformation de /ð^ʕ/ et de l'influence du Berbère, car dans certains dialectes moyen-orientaux on observe ce changement : /ð^ʕ/ > /d^ʕ/ ou /z^ʕ/.

longs siècles du fait qu'elle n'a subi ni de dépeuplements ni d'invasions. Le parler de la ville de Fès fait partie des parlers de sédentaires citadins et de la zone des dialectes fāsī comme décrit par Ennaji (2002_b). Différentes études dialectologiques sur cette ville ont vu le jour (voir les travaux de Lévi-Provençal, 1921* ; Laabi-Zemmama, 1975* ; Rhardisse, 1995 ; Caubet, 1993, 1998* ; Ennaji, 2002_b). D'après Rhardisse (1995), la préservation du parler de Fès est une tâche difficile car la ville a subi différents types de migrations rurales, ainsi que le développement médiatique et l'utilisation imposée de la langue arabe classique. Depuis l'indépendance, le parler de Fès a laissé sa place petit à petit au parler marocain « standard ».

Au niveau consonantique²², le parler de Fès présente des caractéristiques des parlers de sédentaires citadins, à savoir la réalisation sourde du qāf (voir Cantineau (1960), Rhardisse (1995), etc...) et une réalisation occlusive des consonnes interdentes /θ ð ð^s/ ⇒ /t d d^s/. La consonne /t/ est réalisée généralement avec une affrication /t/ ⇒ /tʃ^s/. Cantineau (1960) et Rhardisse (1995), entre autres, expliquent la présence de la paire /r r^s/ ainsi que /z z^s/.

Pour ce qui concerne le système vocalique du parler de Fès, les mêmes schémas sont proposés par la littérature, i.e., un système triangulaire à trois voyelles pleines /i a u/ et une ou deux voyelles centrales brèves /ə/ ou /ə u/, respectivement.

1.5. CONCLUSIONS

Ce chapitre introductif a été principalement consacré aux systèmes vocaliques de l'arabe. Nous avons proposé une réflexion sur ce qu'est exactement la voyelle en arabe et ce à partir des définitions des grammairiens arabes et de sa réalisation, en production et en perception, par des locuteurs natifs. Nous avons développé la situation linguistique complexe de la langue arabe en abordant les questions de son histoire et des liens qu'elle a avec les autres langues **afro-asiatiques**. Nous avons exploré ensuite sa structuration et son organisation en fonction des différents registres, et avons terminé en proposant une typologie

²² Comme pour l'arabe jordanien, nous ne décrivons pas le système consonantique car nous nous intéressons dans ce travail au seul système vocalique. L'arabe marocain de Fès présente un système consonantique différent de celui de l'arabe classique : /b ṭ d k ʔ q m n r f s z ʒ ʃ ʁ h ʕ fi w j l b^s t^s d^s s^s z^s r^s/.

²³ Phénomène présent dans tous les parlers marocains (Cantineau, 1960).

des dialectes arabes modernes du point de vue des classifications géolinguistique et sociolinguistique. Nous avons terminé ce chapitre par la description des deux dialectes étudiés dans cette thèse.

La plupart des études mentionnées traitant du système vocalique des différents dialectes arabes sont basées sur l'étude des voyelles dans des environnements consonantiques et/ou syllabiques précis. Les conclusions tirées de ces études vont dans le sens de systèmes vocaliques triangulaires à 3-5 qualités vocaliques longues et à 1-3 qualités vocaliques brèves ayant un statut phonologique à part entière, en arabe **classique** et en arabe **dialectal**.

Nous avons vu que le système vocalique de l'arabe jordanien est décrit phonologiquement comme composé de cinq timbres vocaliques avec huit voyelles : /i iː eː aː oː uː uː/. Dans des précédentes études (voir Al-Tamimi, 2002 ; Al-Tamimi & Barkat-Defradas, 2002 ; Defradas *et al.*, 2003), nous avons trouvé qu'en production et en perception de la parole, ce système peut être mieux décrit comme /i iː eː ä aː oː uː uː/, en contexte dental. Pour ce qui concerne le système vocalique de l'arabe marocain, les mêmes problèmes se posent au niveau du statut phonologique des voyelles brèves. Est-ce que le système vocalique de l'arabe marocain est composé de 3, de 4, de 5 ou de 6 voyelles ? Quel est le statut phonologique de ces voyelles ? Est-ce que les locuteurs marocains sont conscients de l'évolution de leur système vocalique ?

Pour conclure cette partie, nous avons vu que les systèmes vocaliques de l'arabe **classique**, de l'arabe **standard contemporain** et de l'arabe **dialectal** diffèrent dans leurs organisations et dans leur nombre de segments dans les différents dialectes. Chaque linguiste utilise différents protocoles expérimentaux et différentes méthodes d'analyse qui lui permettent de caractériser au mieux le système vocalique du dialecte étudié. Mais rares sont les études qui décrivent le fonctionnement de ces systèmes dans différentes structures syllabiques et contextes consonantiques.

Nous allons chercher à mieux caractériser les voyelles dans les deux dialectes arabes qui diffèrent comme on l'a vu au niveau vocalique (i.e., différences en nombre de voyelles), rythmiques (les dialectes maghrébins ont un rythme plus saccadé et faisant partie du groupe

des langues accentuelles, tandis que l'arabe jordanien présente un rythme de type syllabique se rapprochant de la complexité syllabique du français), mais également au niveau de leurs lexiques. Nous effectuerons une première comparaison entre ces deux systèmes vocaliques et établirons une seconde comparaison avec le français qui diffère de l'arabe du point de vue de la morphologie de son lexique et au niveau vocalique (voir *supra*).

Les hypothèses que nous avons formulées en introduction générale et visant à expliquer le comportement des locuteurs arabophones vis-à-vis des voyelles produites et perçues en isolation, sont fonction soit de la morphologie du lexique de la langue arabe, soit de la voyelle elle-même (i.e., sa caractérisation du point de vue **statique** et **dynamique**). Nous nous intéressons dans le chapitre suivant à la description de la voyelle en fonction des différents modèles de perception prenant en compte des indices **statiques** et/ou **dynamiques**.

Chapitre 2.

LES MODÈLES DE PERCEPTION DES VOYELLES ET DE LA PAROLE

« *Much of the phonetic information in the signal resides in the consonants.* »

(Miller, 1951).

« *However, a written presentation creates a potential artifact which, if it occurred, could account for the entire advantage of syllable over phoneme targets. If a listener reads the letter « b », he in fact produces (unobserved by the experimenter) a syllable (typically, “bee” or “buh”).* »

(McNeill & Lindig, 1973).

2.1. INTRODUCTION

Nous nous concentrerons dans ce chapitre sur la caractérisation de la voyelle en tant qu'élément **statique** vs **dynamique**. Comme nous le verrons, le débat autour de ces notions alimente les recherches en production et perception de la parole depuis les années 60 et continue à susciter des controverses (voir Strange (1989)). Nous allons nous appuyer sur deux modèles : **statique**, ou l'approche « *Simple Target* » où seule la **cible vocalique** est prise en compte et **dynamique**, ou les approches « *Elaborated Target* » et « *Dynamic Specification* » où les indices **dynamiques (intrinsèques et extrinsèques)** des voyelles sont pris en compte. Nous nous attarderons sur les caractéristiques spécifiques de chaque approche.

Si les caractéristiques **dynamiques** des voyelles jouent un rôle important par rapport aux **statiques**, nous nous attendons à observer plus de variations dans la caractérisation des voyelles avec des éléments **statiques** et plus d'« invariance » avec des éléments **dynamiques**. Les informations **dynamiques** peuvent par ailleurs être complémentaires des informations **statiques**.

Une fois les indices acoustiques pertinents identifiés, nous nous intéresserons aux différentes théories de perception de la parole. Nous développerons ainsi les théories dites « *Passives* » vs « *Actives* ».

2.2. LES APPROCHES EN PERCEPTION DES VOYELLES²⁴

Différentes approches en perception des voyelles cherchant à identifier les indices acoustiques nécessaires à leur caractérisation ont vu le jour, parmi lesquelles : les approches « *Simple Target* », « *Elaborated Target* » et « *Dynamic Specification* ».

2.2.1. L'approche « *Simple Target* »

Généralement, les valeurs moyennes de l'état stable de la voyelle sont considérées comme valeurs de référence permettant sa description et sa caractérisation. Cette vision de la voyelle a été largement admise par les partisans de l'approche « *Simple Target* » pour lesquels les indices **statiques** permettant de décrire une voyelle sont considérés comme les seuls suffisants à leur identification en perception.

Pour les tenants de cette approche, la voyelle est représentée par une forme canonique précise appelée la **cible vocalique**. Elle est définie en tant que représentant d'un segment phonologique indépendamment de son contexte (Joos, 1948 ; Ladefoged, 1967 ; Daniloff & Hammarberg, 1973). Elle est considérée comme un concept unifiant différentes caractéristiques articulatoires, acoustiques et perceptives des voyelles.

Articulatoirement, la **cible vocalique** est représentée par une configuration « stable » du conduit vocal lors de la production d'une voyelle monophthongue en isolation. En parole spontanée, ces positions articulatoirement **statiques** sont considérées comme l'état stable final résultant des coarticulations dans les différents contextes syllabiques (MacNeilage, 1970*).

²⁴ Une grande partie de cette section est inspirée des travaux de Winifred Strange (1989, 1999_a).

Acoustiquement, la **cible vocalique** est représentée par un point dans un espace acoustique multidimensionnel composé des deux ou trois premiers formants oraux. Ces fréquences formantiques sont dérivées à partir d'une seule section spectrale de l'état stable d'une voyelle (Joos, 1948 ; Peterson, 1952, 1961*).

Perceptivement, seuls les deux (ou trois) premiers formants, dans un modèle de perception simple, constituent les informations de base de l'identité perceptuelle d'une **cible vocalique** (Delattre *et al.*, 1952).

La plupart des études défendant cette approche se sont basées sur la perception des voyelles représentées par des stimuli synthétiques soutenus, constitués de patterns de deux formants. Un pattern de deux formants est constitué d'un formant principal et d'un formant secondaire « effectif » nommé F'. Ce formant « effectif », qui est le résultat d'intégration des fréquences de deux formants proches : F_1 & F_2 pour les voyelles postérieures et F_2 & F_3 , pour les antérieures, a une fréquence équivalente à la moyenne des deux pics spectraux proches (Delattre *et al.*, 1952). Ce formant « effectif » joue un rôle important pour expliquer la perception des voyelles.

Pour expliciter davantage le rôle du formant « effectif », nous prenons l'exemple du travail effectué par Chistovich et ses collègues (voir Chistovich, 1985). Leur travail montre que lorsque deux pics spectraux (ou plus de deux pics) se situent à l'intérieur d'une distance critique de 3 à 3,5 Bark de l'échelle fréquentielle psychoacoustique, la qualité de la voyelle perçue est équivalente à un pattern avec un pic spectral unique situé au **centre de gravité** du regroupement des formants. La fréquence formantique « effective » se déplace vers la fréquence du pic d'amplitude le plus élevé. Lorsque la distance fréquentielle entre les pics spectraux dépasse les 3,5 Bark, les formants sont considérés comme perceptivement distincts et les changements de leurs amplitudes relatives n'affecteront pas la qualité des voyelles perçues.

Dans cette approche, les voyelles soutenues sont perçues sur la base de la position de leurs pics spectraux. La plupart des voyelles peuvent alors être discriminées en utilisant les deux premiers formants seulement ; F_1 & F_2 . Ainsi, les voyelles postérieures /u ʊ o ɔ a/ sont

perçues en fonction du **centre de gravité** de F_1 & F_2 , ou F_1' . Les voyelles antérieures /i ɪ e ε æ/ quant à elles, sont perçues en fonction de F_1 & F_2' (F_2' = le **centre de gravité** de F_2 & F_3 , (et même F_4 pour /i/)).

C'est cette représentation de la voyelle, en terme de **cible vocalique** statique, qui a conduit les partisans de cette approche à prendre en compte les valeurs centrales des formants des voyelles comme seules et uniques indices pertinents pour la perception des voyelles par les auditeurs.

Or, même si la **cible vocalique** peut être produite d'une manière quasi stable en isolation, elle doit être considérée comme un « **Artefact de Laboratoire** » (Lieberman *et al.*, 1967_a ; Liberman, 1970), car :

- Les voyelles sont rarement produites en isolation. Elles sont réalisées en coarticulation avec les consonnes adjacentes dans une structure syllabique spécifique (Lieberman *et al.*, 1967_a ; Liberman, 1970),
- La qualité vocalique ne se situe pas à un endroit précis du signal acoustique, mais elle est distribuée sur la totalité de la période de voisement (Harris, 1953 ; Schatz, 1954),
- Les informations acoustiques des voyelles sont confondues avec celles des consonnes (Stevens & House, 1963* ; Lindblom & Studdert-Kennedy, 1967),
- Les formants vocaliques ne sont pas invariants (Peterson & Barney, 1952).

Bien que les paramètres spectraux d'une voyelle isolée ne soient pas invariants, différentes études perceptives ont montré que les auditeurs naïfs étaient capables d'identifier~discriminer les **cibles vocaliques** produites en isolation. Néanmoins, la plupart de ces études ont expliqué que les voyelles présentées en isolation étaient moins bien perçues par rapport à celles présentées en contextes consonantiques (Fairbanks & Grubb, 1961* ; Fujimura & Ochiai, 1963 ; Lehiste & Meltzer, 1973 ; Strange *et al.*, 1976, 1983 ; Strange, 1989, 1999_a ; etc...). La raison avancée est la suivante : puisque la structure acoustique des voyelles produites en isolation est différente de celle des **mêmes** voyelles produites en contexte, les indices utilisés par les auditeurs afin d'identifier ces voyelles (en contexte ou en isolation) doivent être différents.

La recherche s'est alors concentrée sur l'étude de différents indices supplémentaires et **dynamiques** pouvant favoriser la perception des voyelles, (selon les approches « *Elaborated Target* » et « *Dynamic Specification* »). Les tenants de ces approches expliquent que les auditeurs sont sensibles aux informations **dynamiques** propres aux voyelles (i.e., contexte environnant, transitions formantiques) et supplémentaires (comme f_0 et durée, formants supérieurs, etc...) en complément des indices **statiques**.

2.2.2. L'approche « *Elaborated Target* »

D'après cette approche, différents indices « élaborés » peuvent coexister pour permettre aux auditeurs une meilleure compréhension des voyelles, à savoir des indices **intrinsèques** (f_0 et plus particulièrement le rapport F_1/f_0 , le rapport entre les formants, les formants supérieurs, etc...) et **extrinsèques** (les patterns formantiques en fonction du contexte environnant et en fonction de l'organisation globale de l'espace vocalique pour un locuteur donné, etc...). Cette approche a été développée essentiellement pour la prise en compte de la variabilité inter-individuelle entre des locuteurs de sexes et âges différents.

Afin d'illustrer les caractéristiques principales de cette approche, nous prenons l'exemple classique de l'étude proposée par Peterson & Barney (1952). Les auteurs ont étudié la distribution des voyelles de l'anglais américain produites par 76 locuteurs de différents sexes et âges (33 hommes, 28 femmes et 15 enfants). Les dix voyelles monophthongues (/i ɪ ε æ ʌ ɑ ɔ u ʊ ə/) de l'anglais américain sont produites dans le contexte [hVd]. Le but de cette étude est de montrer que la dispersion acoustique des voyelles en anglais américain est influencée par les différentes réalisations individuelles. Leurs résultats indiquent que pour chacun des segments vocaliques, une variabilité inter-individuelle importante est observée (voir Figure 9). Elle est imputée aux différences physiologiques et dialectales : les hommes parlent l'anglais américain Standard, tandis que les femmes et les enfants parlent l'anglais américain de la zone Atlantique.

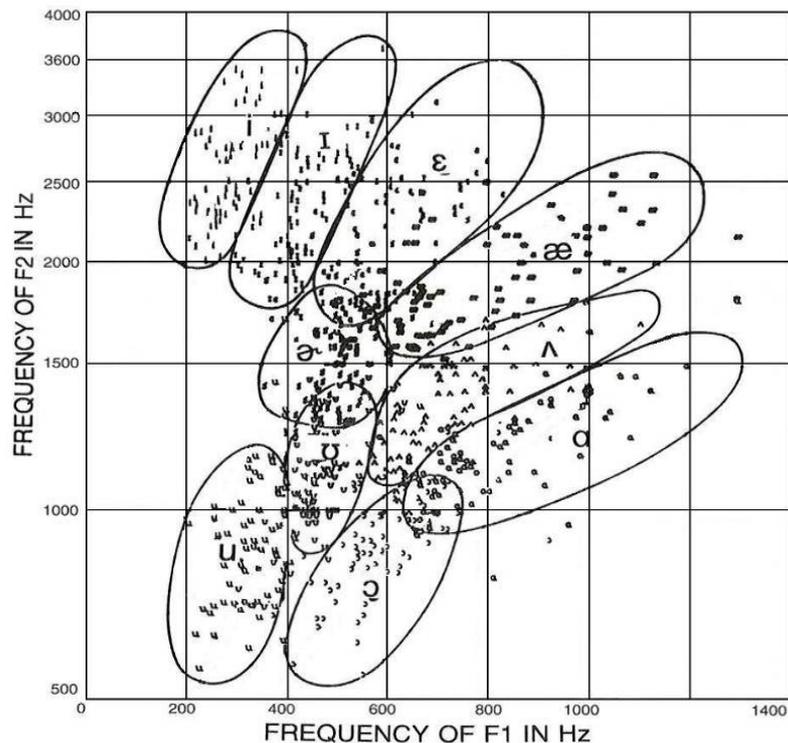


Figure 9 : Espace vocalique représentant la dispersion des 10 voyelles de l'Anglais Américain produites par 76 locuteurs. On observe une grande variabilité due aux locuteurs. (Emprunté à Peterson & Barney, 1952).

La question qui peut se poser est alors de savoir : comment les auditeurs arrivent-ils à identifier les voyelles produites par les hommes, les femmes et les enfants, malgré la grande variabilité observée dans les fréquences formantiques « absolues » ?

Différents auteurs ont montré que l'information contenue dans f_0 (et surtout le rapport F_1/f_0), le rapport entre les formants et les formants supérieurs (spécialement F_3) apportent plus de spécification quant à la qualité de la voyelle. (Voir les travaux de Ryalls & Lieberman, 1982, sur f_0 et le rapport F_1/f_0 ; Syrdal & Gopal, 1986, sur l'utilisation d'une représentation bidimensionnelle avec les valeurs de F_1-f_0 (en Bark) représentées en abscisse et celles de F_3-F_2 en ordonnée, qui semble réduire la variabilité observée dans les données de Peterson & Barney, 1952, etc...).

En ce qui concerne le rôle des rapports entre les formants des voyelles, James Miller (voir Miller, 1989, pour une revue de littérature de cette théorie) a développé « *The Formant-Ratio Theory of Vowel Perception* », qui répond à la question de la normalisation des valeurs formantiques des voyelles produites par des locuteurs différant par l'âge et le sexe. Cette théorie propose que l'identité d'une voyelle dépende des intervalles entre les formants, le ratio

formantique ou le rapport entre les formants, plutôt que des valeurs absolues des fréquences de formants. Pour la voyelle /i/, la distance entre F_1 et F_2 est grande, résultant ainsi en un ratio important. En revanche, pour la voyelle /a/, la distance entre F_1 et F_2 est basse, ainsi le ratio est bas. La Figure 10 montre la dispersion des voyelles de l'anglais américain à partir des données de Peterson & Barney (1952) et les courbes et les ratios des rapports de formants pour les voyelles correspondantes.

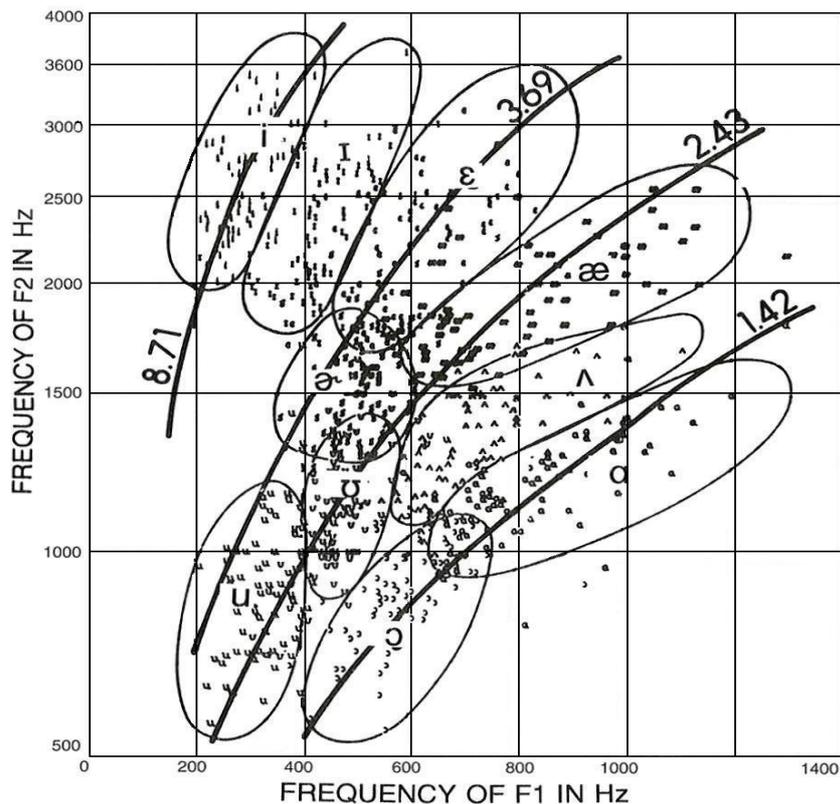


Figure 10 : Triangle vocalique représentant les fréquences F_1 & F_2 des 10 voyelles de l'anglais américain produites par 76 locuteurs (*hommes, femmes* et enfants). Les courbes représentent les ratios constants de F_2/F_1 en fonction des voyelles. (Emprunté à Miller, 1989).

Les rapports sont : 8,71 pour /i/, 3,69 pour /ε ε u/, 2,43 pour /æ u u/ et 1,42 pour /a ɔ/. Ces ratios semblent être utilisés par les auditeurs pour compenser les différences de taille de conduit vocal entre les hommes, les femmes et les enfants. Le problème de ces rapports est le fait que des voyelles acoustiquement différentes obtiennent le même ratio de F_2/F_1 , comme c'est le cas de /æ u u/ où le ratio est de 1,42 (la voyelle /u/ obtient 2 ratios = 1,42 et 2,43 dans ce modèle). Dans ces conditions, comment ces voyelles à fréquences absolues différentes et à ratios identiques peuvent-elles être discriminées par les auditeurs ? Différentes études ont

montré que les auditeurs utilisent d'autres indices que le rapport de formant, tels que f_0 (ou le rapport entre F_1/f_0) (voir *supra*), mais également l'utilisation des formants supérieurs.

En effet, Fujisaki & Kawashima (1968*) montrent que le changement des valeurs de F_3 affecte la perception des voyelles synthétiques du japonais et plus particulièrement lorsque ces modifications sont accompagnées de changement de valeurs de f_0 . Pour Lehiste & Meltzer (1973), l'utilisation des valeurs de f_0 et des fréquences de formants (F_1 , F_2 et F_3) propres aux hommes, aux femmes et aux enfants aident les auditeurs à mieux discriminer les voyelles produites avec des différences d'âge et de sexe. Toutefois, leurs résultats ne montrent pas systématiquement une préférence pour le rapport entre F_1 et f_0 pour la discrimination des locuteurs selon le sexe et l'âge.

Par ailleurs, Miller (1989) montre qu'outre les ratios de formants vocaliques utilisés par les auditeurs, la représentation des voyelles dans un espace perceptivo-auditif ajoute une séparation des voyelles et permet ainsi une meilleure discrimination entre celles-ci. Sa méthode est basée sur la transformation des valeurs absolues de formants, des hommes, femmes et enfants, en unités de fréquence logarithmique et de calcul des intervalles²⁵ de celles-ci. Les intervalles correspondent aux valeurs de F_3-F_2 , F_2-F_1 et de F_1-SR (où SR est une « référence sensorielle » dérivée de f_0 du locuteur). Les résultats de ces études montrent une réduction des chevauchements observés dans la représentation traditionnelle des voyelles obtenues dans l'étude classique de Peterson & Barney (1952), (voir Figure 11).

En ce qui concerne l'utilisation des formants supérieurs (i.e., F_3 , F_4 & F_5), l'étude de Nearey (1989) a mis en évidence leur rôle dans la catégorisation des voyelles de l'anglais américain. Les résultats qu'il a obtenus montrent que la catégorisation des voyelles change en fonction à la fois de l'augmentation ou de la baisse des valeurs des formants supérieurs et des valeurs de f_0 .

Outre l'utilisation des informations **intrinsèques** propres aux voyelles, différentes études ont contribué à la caractérisation du rôle des informations **extrinsèques**, qui affectent

²⁵ Un intervalle d'une unité logarithmique = le ratio des unités linéaires.

la totalité de l'espace vocalique d'un locuteur, pour une meilleure discrimination des voyelles. Différentes méthodes de normalisation du locuteur sont basées sur ces informations **extrinsèques**.

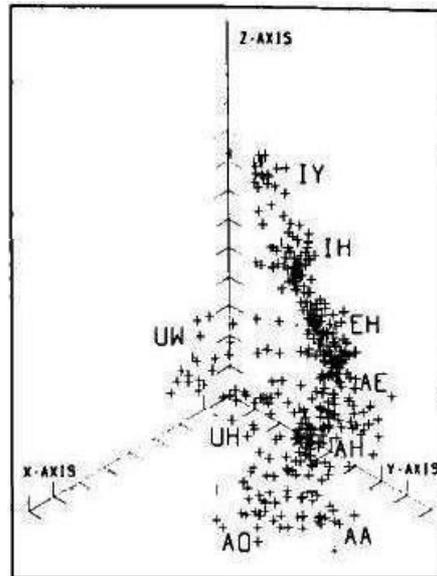


Figure 11 : Espace perceptivo-auditif de 406 voyelles non-rétroflexes de l'anglais américain des données de Peterson & Barney (1952). L'axe X correspond aux valeurs de F_3-F_2 , l'axe Z aux valeurs de F_2-F_1 et l'axe Y aux valeurs de F_1-f_0 . (Emprunté à Miller, 1989).

D'après ces approches, les auditeurs interprètent les patterns de formants des voyelles en référence au contexte environnant (i.e., le contexte consonantique et/ou syllabique, etc...), mais également en fonction de l'ensemble des patterns formantiques de l'espace vocalique d'un locuteur, ou les informations structurelles (Joos, 1948). Même si les valeurs formantiques absolues des voyelles produites par différents locuteurs de différents sexes et âges varient largement, la relation entre les patterns formantiques de ces voyelles est identique entre les locuteurs. La voyelle /i/ a un F_1 très bas par rapport à /ɪ/ et un F_2 très haut par rapport à toutes les voyelles à travers les locuteurs. La voyelle /u/ présente un F_1 et un F_2 très bas et la fréquence de F_1 pour /a/ est la plus élevée. L'auditeur percevant les voyelles de n'importe quel locuteur, va se situer dans un espace vocalique délimité par les voyelles extrêmes /i a u/ produites par le locuteur. Ensuite, c'est une question de pattern de formants relatif qui va guider l'auditeur pour l'identification des autres voyelles (Ladefoged & Broadbent, 1957 ; Strange, 1989 ; etc...).

Il semblerait que les auditeurs soient sensibles aux patterns formantiques des hommes, femmes et enfants. En effet, comme ces informations sont liées à la longueur du conduit vocal du locuteur²⁶, l'auditeur est capable de détecter ces différences entre eux. Néanmoins, des études ont montré que les auditeurs ont un taux d'erreur élevé dans une tâche d'identification des voyelles produites par un même locuteur lors de différentes sessions, « *Mixed Speaker test* » en comparaison avec celles produites en une seule session, « *Blocked Speaker Test* » (voir Strange *et al.*, 1976 ; Macchi, 1980 ; Assmann *et al.*, 1982 ; etc...). Toutefois, Verbrugge *et al.* (1976) ont montré que lorsque les trois voyelles extrêmes /i a u/, produites par le même locuteur, sont présentées avant les stimuli, dans la même tâche d'identification de voyelles (i.e., *Mixed Speaker test*), les auditeurs identifient mieux les voyelles de chaque locuteur (voir les travaux de Macchi, 1980 ; Assmann *et al.*, 1982 ; etc..., pour des résultats inverses).

Ces différentes études ont mis en évidence le rôle de différents indices **intrinsèques** (f_0 et les formants supérieurs) et/ou **extrinsèques** (les gammes formantiques de chaque locuteur de différents sexes et âges) dans l'identification des voyelles synthétiques et naturelles produites dans différentes syllabes. Le fait que les auditeurs identifient des voyelles naturelles mieux que les voyelles synthétisées montrent que l'information syllabique contient suffisamment d'indices sur la qualité de la voyelle (i.e., les transitions formantiques, les Changements Spectraux Inhérents aux Voyelles, etc...). Ce sont ces derniers qui ont été mis en avant par les tenants de la troisième approche de perception des voyelles, l'approche « *Dynamic Specification* ».

2.2.3. L'approche « *Dynamic Specification* »

Ce modèle de perception des voyelles vise à prendre en compte les aspects **dynamiques** entourant la voyelle. Différents auteurs ont ainsi expliqué les différences entre les voyelles produites en isolation ou en contexte [hVd] par rapport à celles produites en

²⁶ En moyenne, un homme a un conduit vocal de 17,50 cm, celui d'une femme est inférieur de 15% et celui d'un enfant de 50%. Ce qui donne un ratio de réduction de la taille du conduit vocal de 5/6 pour la femme et 1/2 pour l'enfant.

contextes consonantiques ou syllabiques différents (voir par exemple les travaux de Stevens & .

Dans leur étude, Stevens & House (1963*) ont comparé les valeurs formantiques des voyelles produites en isolation, en contexte [hVd] et en contextes consonantiques²⁷. Le corpus est basé sur la production des 8 voyelles /i ɪ ε æ ʌ ʊ u/ en isolation, en contexte [hVd] et en association avec l'une des consonnes /p b f v θ ð s z t d tʃ dʒ k g/ en contexte CVC symétrique. Les valeurs des trois premiers formants ont été mesurées au milieu temporel de chaque voyelle. La Figure 12 montre les effets de contextes consonantiques sur les voyelles.

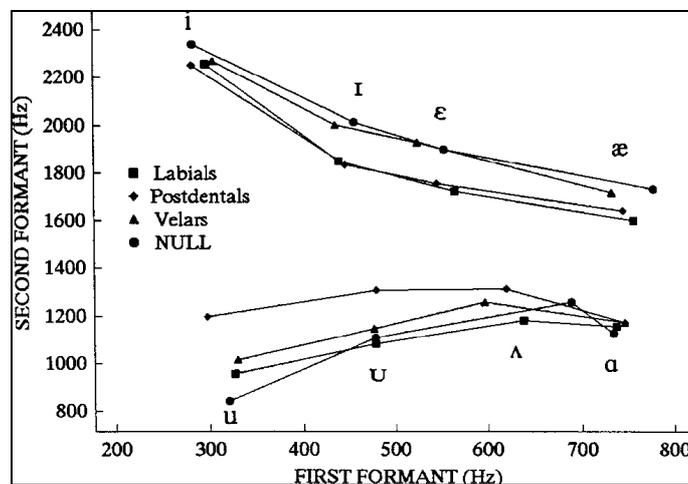


Figure 12 : Schéma montrant les effets des contextes consonantiques et des environnements « nul » sur les voyelles de l'anglais américain. (Emprunté à Stevens & House, 1963*).

Les résultats obtenus montrent un effet de centralisation des voyelles en contexte non « nul » (surtout sur l'axe F_2 , en moyenne de ± 200 Hz selon la voyelle), ainsi qu'un abaissement des valeurs de F_1 des voyelles en contexte consonantique voisé. Le mode d'articulation affecte également les voyelles : les consonnes occlusives augmentent les valeurs de F_2 par rapport aux autres modes d'articulation.

Pour Stevens & House (1963*), la différence observée entre les valeurs formantiques des voyelles en isolation ou en [hVd] en comparaison avec le contexte consonantique s'explique par un modèle de réduction vocalique (Lindblom, 1963_a). Pour les auteurs, le

²⁷ Les auteurs nomment les contextes de réalisations de voyelles en isolation ou en [hVd] comme contexte « nul » n'ayant pas d'influence sur la qualité de la voyelle.

système de production des voyelles est dérivé d'un modèle de **cible vocalique** correspondant à des postures articulatoires précises dans un environnement nul, tandis qu'en contexte consonantique, la **cible vocalique** idéale n'est pas réalisée à cause de contraintes d'inertie. Ils suggèrent que les auditeurs utilisent leurs connaissances des effets consonantiques afin de reconnaître les voyelles :

« The rules governing these deviations in the acoustic signal must, of course, be invoked in some way by the listener in order to make an identification of the signal ».

(Stevens & House, 1963* : p. 122)

Les influences consonantiques sur les voyelles sont principalement dues au phénomène de coarticulation²⁸. En effet, en isolation, la réalisation d'une voyelle n'est influencée par aucun obstacle dans le conduit vocal. Dès lors qu'on demande au locuteur de réaliser une voyelle en contexte, un changement de forme du conduit vocal de cette voyelle, selon la consonne qui précède et/ou qui suit est observé (voir Figure 13). Dans cet exemple classique, un locuteur français produit la voyelle /u/ en isolation (en pointillé sur le graphique) et en contexte dental comme dans la réalisation du mot : doute > /dut/ (en ligne discontinue). Nous remarquons que la production de la voyelle en contexte CVC est influencée par le lieu d'articulation, provoquant ainsi une antériorisation de la production de la voyelle postérieure. Au niveau phonétique, la réalisation de cette voyelle postérieure antériorisée est proche de la réalisation de [y].

²⁸ D'autres facteurs peuvent expliquer la modification des valeurs formantiques des voyelles produites en isolation et en contexte, comme : le débit et styles de parole, l'état émotionnel du locuteur, les différences physiologiques, sociales et régionales, que nous ne développerons pas ici. (Voir Pickett, 1999, pour le débit et styles de parole et les différences physiologiques ; Caelen-Haumont, 2005, pour l'état émotionnel du locuteur ; Chambers *et al.*, 2002, pour les différences sociales et régionales).

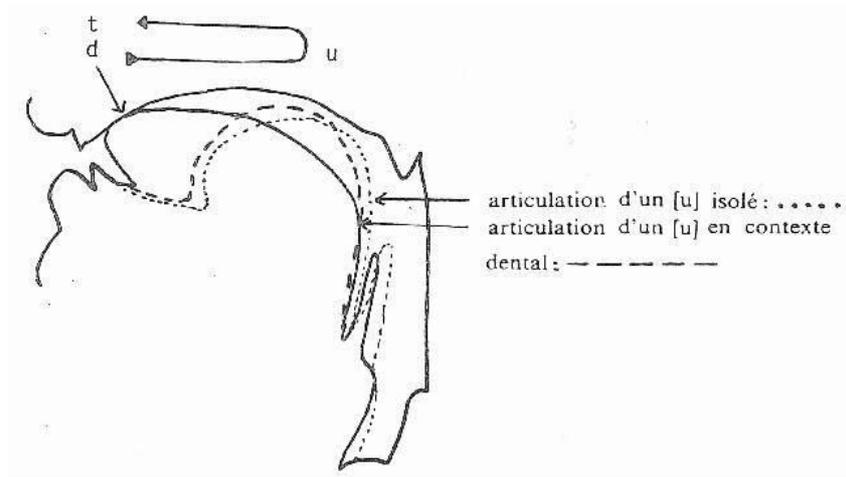


Figure 13 : Déplacement des articulateurs en contexte antérieur (dental) vs postérieur lors de la réalisation de la voyelle /u/ en français dans le mot /dut/. (Emprunté à Calliope, 1989).

Nous savons que ces voyelles subissent une forme de réduction vocalique nommée « *Target Undershoot* » (Lindblom, 1963_a). Cette réduction vocalique est essentiellement liée à la non-accentuation des voyelles et au débit de parole rapide, ainsi qu'aux effets contextuels, dont les résultats se traduisent par une tendance à la « centralisation » de toutes les voyelles (surtout de types périphériques) vers un [ə] (Joos, 1948 ; Lindblom, 1963_a).

Ces effets contextuels affectent largement les voyelles en modifiant les valeurs formantiques de l'état stable et ainsi leur qualité. Ces modifications ne se limitent pas à l'état stable, mais affectent largement la forme des transitions et des trajectoires inhérentes aux voyelles mêmes. Ces modifications de la « couleur » des voyelles affectent également l'identification de celles-ci par les auditeurs qui semblent utiliser différents indices acoustiques, outre les valeurs de formants et les rapports existant entre eux.

Ces différents indices sont considérés ainsi comme **dynamique**, i.e., prenant en compte les trajectoires de formants dûes aux facteurs **intrinsèque** (les trajectoires **dynamiques** et les Changements Spectraux Inhérents aux Voyelles) et **extrinsèque** (les transitions formantiques liées aux contextes consonantiques et/ou vocaliques environnant).

En ce qui concerne les indices **dynamiques intrinsèques** liés aux Changements Spectraux Inhérents aux Voyelles (CSIV), les travaux sont nombreux (voir Nearey & Assmann, 1986 ; Hillenbrand *et al.*, 1995 ; Nearey, 1989 ; Hillenbrand *et al.*, 2001 ; etc...). Le principe de base de cette méthode est de caractériser la voyelle à partir de deux valeurs

formantiques (au début et à la fin de la voyelle) différentes de celle obtenue au milieu temporel de la voyelle (à 50%). Dès lors, cette méthode propose de quantifier l'étendu des trajectoires des formants liées à la voyelle même en gommant les influences consonantiques présentes dans les transitions). Cette méthode évalue des trajectoires formantiques systématiques propres aux voyelles (comme pour les voyelles monophthongues de l'anglais américain ou canadien).

Nous prenons comme exemple l'étude de Nearey & Assmann (1986), qui ont montré que les auditeurs étaient capables d'identifier les voyelles produites en isolation à partir des fréquences formantiques extraites à 30 ms après l'Onset et 30 ms avant l'Offset. Hillenbrand *et al.* (1995) ont montré que les taux de classification correcte de l'Analyse Discriminante²⁹ étaient de 88,8% en utilisant les valeurs des CSIV à 20% et à 70% et de la durée vocalique (en comparaison avec les taux obtenus en utilisant la seule valeur à 50%, avec une amélioration des taux de 4%). Dans la même optique, Hillenbrand *et al.* (2001) ont montré que le taux de classification correcte de l'Analyse Discriminante a été amélioré de 6% avec les valeurs CSIV obtenues 20% et à 80% en comparaison avec la valeur de l'état stable obtenue à 60% (92% avec les CSIV et 86% avec la valeur de l'état stable).

En ce qui concerne les transitions vers et à partir des voyelles, l'étude réalisée par DiBenedetto (1989_a, 1989_b) montre que la trajectoire de F₁ influence largement la perception des voyelles /i i e ε æ/, dans une structure de type [dVd]. L'auteur a noté des variations dans l'identification de la hauteur des voyelles et dans leur classification comme voyelles tendues ou relâchées, en fonction de la durée relative de l'Onset et de l'Offset de F₁. DiBenedetto a conclu en expliquant que la qualité des voyelles dépend des paramètres temporels liés à la totalité de trajectoire de F₁.

²⁹ L'Analyse Discriminante peut être définie comme une des techniques de « scoring » pour déterminer la contribution des différentes variables qui expliquent l'appartenance des individus à des groupes. Deux ou plusieurs groupes sont comparés, sur plusieurs variables pour déterminer s'ils diffèrent et pour comprendre la nature de ces différences. Nous reviendrons plus en détails sur comment cette analyse fonctionne, car nous l'utiliserons dans notre thèse.

Pour ce qui concerne les indices **dynamiques extrinsèques** aux voyelles, l'étude sur le « *Target Undershoot* » établie par Björn Lindblom (Lindblom, 1963_a), montre que la réduction vocalique due au contexte consonantique dans un environnement CVC est compensée par les auditeurs qui tentent de rétablir les **cibles vocaliques** canoniques. Cette compensation est effectuée par « *Perceptual Overshoot* » (Lindblom & Studdert-Kennedy, 1967), où la durée vocalique réduite est utilisée par les auditeurs pour compenser cette réduction. Pour les auteurs, ce mécanisme de « *Perceptual Overshoot* » naît à partir d'un processus d'adaptation que l'auditeur opère en fonction de la réduction vocalique, causée par la durée réduite et les changements de fréquences formantiques :

« During the test stimulus, auditory events can be grossly pictured as follows: There occurs a continuous change in the sensitivity of the ear owing to adaptation. At any given moment the "tracking" of the formant peak may be shifted towards a region of lower threshold, that is, upwards for an ascending ramp and downwards for descending ramps, the amount of displacement depending on how far adaptation has progressed at that moment. This effect would accordingly be one of the factors contributing to shifting matches closer to the terminal frequencies of the ramps. If verified by future experimentation these speculations invite the conclusion that the peripheral auditory representation of spectral formant peaks depends on the rate of formant frequency change and offer a possible explanation of some of the present rate- and context-dependent boundary shifts. According to this view, vowel production and vowel perception could be said to be complementary in the sense that articulatory activity is characterized by undershoot and perception by overshoot. »

(Lindblom & Studdert-Kennedy, 1967 : p.841)

Ces résultats montrent que la perception des voyelles coarticulées dans une structure CVC n'est pas basée exclusivement sur les valeurs centrales des formants, mais davantage sur la direction et la pente des transitions formantiques. Lindblom (1963_a) a proposé une équation permettant de mesurer la compensation perceptive. Cette formule est dérivée : 1) des valeurs de l'Onset et de l'Offset (valeurs de « locus virtuel », Delattre *et al.*, 1955 ; voir § 4.3.2.1.1) ; 2) des valeurs des cibles vocaliques ; et 3) des durées des syllabes. Cette proposition de compensation perceptive repose essentiellement sur l'idée selon laquelle la réduction vocalique articulatoire et acoustique augmente systématiquement lorsque la durée de la syllabe diminue, et vice-versa. Différentes études ont cependant mis en cause l'effet direct de la durée des syllabes sur la réduction vocalique. Gay (1978) ; Kuehn & Moll (1976) ; Pols &

van Son (1993) ; et Moon & Lindblom (1994) ; entre autres, ont montré que la réduction vocalique ne dépend pas uniquement de la durée mais également des locuteurs ou du débit de parole. Certains locuteurs compensent la réduction de la durée des syllabes soit par une accélération de leurs gestes articulatoires, soit par une production des gestes vocaliques pendant (ou même avant) la réalisation du geste consonantique (Gay, 1977), et atteignent par conséquent les mêmes cibles vocaliques en débit lent et rapide. Gay (1978) et Pols & van Son (1993) ont montré que la durée n'affecte pas la réduction vocalique en comparant des voyelles accentuées produites à débit rapide et non-accentuées produites à débit lent.

Comme nous l'avons rappelé, l'identification des voyelles coarticulées en contexte CVC obtient des scores plus élevés que l'identification des voyelles produites en isolation (voir Strange *et al.*, 1976, 1979*, 1983; Gottfried & Strange, 1980; Strange, 1989, 1999_a; Case *et al.*, 1995; Jenkins & Strange, 1999; etc...). La plupart de ces études ont mis l'accent sur la meilleure identification des voyelles en contexte par rapport aux voyelles isolées. À l'inverse, certains auteurs ont montré que le contexte consonantique environnant n'apportait pas plus de précision quant à la qualité de la voyelle par rapport à celle en isolation (Macchi, 1980 ; Assmann *et al.*, 1982 ; Jenkins *et al.*, 1999 ; Pitermann, 2000 ; etc...). Ces différences peuvent être imputées soit à la tâche des auditeurs (i.e., identification des voyelles en fonction des syllabes de type [pVp] écrite orthographiquement, qualité des stimuli, etc...), soit aux indices utilisés dans l'identification des voyelles (i.e., durée de transitions et des voyelles, centres syllabiques recombinaés, etc...). C'est pourquoi, nous proposons dans ce travail d'évaluer le rôle des indices **statiques** et **dynamiques** en proposant le **même** type de stimuli (i.e., stimuli synthétiques à durée égale) et la même tâche (i.e., identification des prototypes vocaliques et catégorisation des espaces vocaliques).

Pour évaluer le rôle des transitions formantiques (à l'**onset** et à l'**offset**) en perception des voyelles, Winifred Strange et ses collègues ont élaboré une méthode basée sur l'extraction des voyelles de leurs contextes CVC.

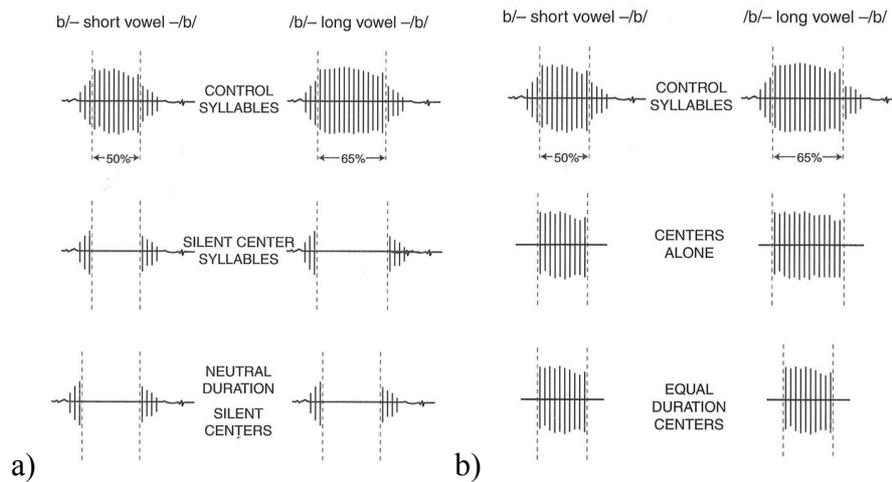


Figure 14 : Schémas représentant la méthode d'extraction utilisée par Winifred Strange et ses collègues pour l'étude de l'effet des informations contenues dans les **onsets** et **offsets** (a) et au centre des voyelles (b), avec et sans normalisation de durée. (Empruntés à Strange, 1999_a).

Les deux schémas de la Figure 14 montrent l'exemple des voyelles longues et brèves produites dans un contexte [bVb]. À gauche (a), les deux voyelles sont extraites et les informations contenues dans les **onsets** et **offsets** sont conservées. À droite (b) en revanche, seule l'information contenue au centre de la voyelle qui est conservée. Une fois ces deux parties séparées, le temps est compressé pour procéder à une normalisation des voyelles longues pour qu'elles aient une durée identique à celle des voyelles brèves.

Leurs résultats sur le rôle des centres syllabiques vides, les *Silent Center Syllables*, montrent que les auditeurs utilisent les informations contenues dans les trajectoires formantiques pour identifier les voyelles coarticulées avec un taux d'identification élevé (Jenkins *et al.*, 1983 ; Strange *et al.*, 1983 ; Parker & Diehl, 1984 ; Strange, 1999_a ; etc...). Alors que l'identification des voyelles sur la base des informations contenues dans l'état stable des voyelles (*Centers Alone*) montre une augmentation du taux d'erreur (voir Strange *et al.*, 1976, 1983 ; Jenkins *et al.*, 1983 ; etc...).

Ces résultats vont dans le sens de la spécification **dynamique** des voyelles en fonction de leurs transitions et non uniquement des centres des syllabes (i.e., **cible vocalique**). Il apparaît également que les informations sur les trajectoires formantiques sont complémentaires et distinctes de celles des états stables des voyelles (lesquelles subissent largement les effets coarticulatoires des consonnes environnantes) :

Ces différentes études montrent l'importance des indices **dynamiques intrinsèques** (les Changements Spectraux Inhérents aux Voyelles) et **extrinsèques** (les transitions formantiques et les trajectoires vocaliques liées aux contextes consonantiques et/ou vocaliques environnant). Ces différents indices permettent une meilleure identification des voyelles coarticulées et/ou en isolation. Ceci a amené différents auteurs à considérer les voyelles comme des événements articulatoires et acoustiques **dynamiques**. Par exemple, Fowler (1980) considère les voyelles de l'anglais américain comme des gestes ayant des caractéristiques temporelles **intrinsèques**.

Les différentes études que nous avons présentées, et qui défendent chacune une des trois approches de perception des voyelles, montrent bien comment l'auditeur peut utiliser et interpréter les différents indices acoustiques **statiques** et/ou **dynamiques** en perception des voyelles. Nous nous intéressons dans les parties qui suivent aux différentes théories de perception de la parole : les « *Théories Actives* » et les « *Théories Passives* », la différence se situe au niveau du rôle de l'auditeur en perception de la parole. Pour les premières, il a un rôle primordial dans la perception des catégories, tandis que pour les secondes, il a un rôle passif car la perception de la parole fait partie intégrante de la faculté de perception générale.

2.3. THÉORIES DE PERCEPTION DE LA PAROLE

La perception de la parole³⁰ est considérée comme un processus de traitement et d'interprétation de la structure imprimée dans le signal acoustique produite par le locuteur, ou comme l'explique Berry (1969*) : « *Auditory perception of speech per se deals mainly with the temporal management of information from the input* », (Berry, 1969* : pp. 59).

La parole est un processus continu, difficile à segmenter dans le temps. Les différents articulateurs évoluent continuellement mais produisent un signal perçu comme une suite de sons constituant des « unités distinctives ». Daniloff (1973*) explique que :

³⁰ Ce chapitre est inspiré en grande partie des travaux suivants : Sanders (1977), Rosner & Pickering (1994), Hawkins (1999_a, 1999_b, 1999_c), tutoriel des modèles de perception de la parole : http://www.unige.ch/fapse/PSY/persons/frauenfelder/SP/Model_speech.html (Fougeron & Cutugno, 1999), Pisoni & Remez (2005) et Tatham & Morton (2006).

« *Articulation processes run together and overlap the phoneme-sized units so that the phonemes are allophonically varied and changeable. From this shifting continuous stream, listeners can decode speech and extract the phonemes units, even though it is difficult to observe independent, non-overlapping sounds in the speech acoustic wave (Liberman et al., 1967). It is this seeming divergence between production and perception (Kozhevnikov et Christovich, 1965) and between larger, perhaps syllable sized production units and smaller perceptual ones which puzzles students of articulation* ».

(Daniloff, 1973* : p. 208).

On note également que la production des sons de la parole nécessite un degré important d'invariance pour que ces sons soient facilement identifiables et compris par l'auditeur. La difficulté réside alors dans la détermination de la nature de cette invariance.

Pour savoir comment les valeurs linguistiques sont obtenues à partir du signal acoustique, deux approches différentes ont été considérées. MacKay (1956*) propose les termes de « *Théories Passives* » et de « *Théories Actives* » pour désigner ces deux approches³¹.

2.3.1. *Théories Passives*

Les « *Théories Passives* » sont présentées par MacKay (1956*) comme des théories basées sur un système de filtrage fonctionnant pour identifier et permettre de combiner différentes informations sur les sons de la parole afin de structurer le pattern perceptif. En effet, ces théories mettent l'accent sur un mécanisme sensoriel et de filtrage ainsi que sur le mappage direct du message vers des traits phonético-acoustiques de la langue. Les partisans de ces « *Théories Passives* » minimisent le rôle du mécanisme de production et de l'articulation des sons du langage : ils lui donnent un rôle mineur et secondaire non utile en perception de la parole (bien qu'il puisse être utilisé dans des circonstances précises). Pour ces théories, l'auditeur n'utilise aucune représentation interne de la production des sons de la

³¹ Outre les théories « *Actives* » et « *Passives* », il existe deux autres catégories générales de théories de perception de la parole. D'une part, on parle de théories « *Bottom-Up* » et « *Top-Down* » qui expliquent la direction de traitement de l'information et la nécessité (ou non) d'avoir toute l'information requise pour identifier les sons dans le signal. De l'autre, on parle des théories « *Autonomes* » et « *Interactives* » pour désigner toute l'information incluse dans le signal acoustique (pour plus de détails, voir Sanders, 1977).

parole, mais il dérive tout simplement l'information linguistique du signal acoustique dans des processus en série. Ces séries sont représentées par les traits distinctifs par exemple, où un son est représenté par une multitude d'informations quant à la sonorité, la nasalité, etc... L'auditeur fait appel à ces diverses informations pour identifier les sons extraits du signal acoustique.

Parmi les différentes « *Théories Passives* », nous avons, le *Modèle Pandémonium*, qui est un modèle computationnel basé sur la détection des traits en parallèle, proposé par Selfridge (1959*). Ce modèle présente les traits distinctifs comme unité primitive de perception de la parole et défend l'existence de l'invariant acoustique puisque tous les phonèmes sont identifiés en fonction des traits qui les composent.

Le deuxième modèle est le *Modèle d'Audition* proposé par Gunnar Fant (Fant, 1960, 1962*, 1967*), qui est un modèle acoustique s'inspirant de la *Théorie Acoustique de la Parole*. Une fois les différents patterns acoustiques d'un phonème identifiés, ils sont transformés en structures neurophysiologiques précises du système auditif. Pour Gunnar Fant, il est possible de proposer un algorithme de reconnaissance de la parole à partir d'une matrice binaire des traits distinctifs liée directement au pattern acoustique transformé. Ces patterns acoustiques sont alors directement et passivement encodés pour déterminer leur appartenance à tel ou tel trait.

Le troisième modèle est le *Modèle de Renforcement Auditif*, développé par l'équipe de Randy Diehl et ses collègues (voir Diehl *et al.*, 1990* ; pour un résumé de l'approche) et utilisé par Kenneth Stevens (voir Stevens *et al.*, 1986 ; Stevens & Keyser, 1989). Ce modèle a pour but l'étude des traits distinctifs redondants des unités phonologiques. Ces traits redondants, une fois combinés, soit augmentent le taux d'identification d'un trait distinctif précis à cause de la présence de celui-ci dans le signal, soit augmentent indirectement les chances de ce trait d'être identifié.

Le quatrième modèle est le *Modèle Neurologique*, développé par Abbs et Sussman (1971*) et Uttley (1959*, 1966*). Il est basé sur un modèle computationnel en réseaux de neurones. Les auteurs supposent l'existence, en perception de la parole, de groupes de cellules

neurologiques capables d'identifier la présence ou l'absence de certains traits distinctifs dans le signal acoustique.

Le cinquième et dernier modèle passif de perception de la parole, a été initialement développé par Gibson (1954*, 1966*) sous le nom *Direct Perception* et utilisé ensuite par Fowler (Fowler, 1986, 1996 ; Fowler & Rosenblum, 1991*) sous le nom *Direct Realism*. D'après ce modèle, la perception de la parole obéit aux mêmes principes que les autres types de perception sensorielle. La perception de la parole utilise des indices structurels intégrés directement dans le signal acoustique (i.e., ils ne sont pas dérivés, ni interprétés à partir du signal). Le geste articulatoire~phonologique est considéré comme une primitive phonologique permettant la structuration du signal acoustique détecté par les organes de perception sensorielle. Pour ce modèle, lorsqu'on perçoit des phonèmes, on perçoit directement les gestes les structurant phonologiquement : les gestes, qui sont discrets et qui se chevauchent, sont ce qu'on produit et ce qu'on perçoit.

La vision apportée par les « *Théories Passives* » de perception de la parole n'est pas admise par tout le monde car, lorsqu'on parle, les variations existantes dans le signal acoustique sont la résultante de la variabilité inter-individuelle. Le rôle de l'auditeur est d'être capable d'identifier ces variations dans le signal acoustique afin de comprendre le message transmis par le locuteur. Autrement dit, le couple locuteur~auditeur a un rôle important dans le mécanisme de perception de la parole. C'est l'idée de base que défendent les partisans de la seconde approche.

2.3.2. *Théories Actives*

À l'inverse des « *Théories Passives* », les « *Théories Actives* » de perception de la parole sont des systèmes de comparaison entre le signal acoustique et un pattern interne (MacKay, 1956*). Ce pattern interne correspond à l'articulation, à la production du signal de la parole. Ces théories de perception sont présentées comme dépendantes du processus de production de la parole où le locuteur~auditeur joue un rôle actif important dans les deux processus de production et de perception de la parole. Les partisans de ces théories postulent

que l'intégration de certains aspects de la production de la parole dans les processus de perception permet à l'auditeur d'être conscient de ce qu'il entend, mais également d'être capable d'analyser les sons de la parole par rapport à leurs propriétés phonétiques en référence à la manière dont ils sont produits. Ils sont ensuite reconnus par l'auditeur, sur la base des correspondances établies entre les unités perceptives et les représentations générées à partir des patterns internes de la production des sons. Ces patterns internes peuvent être le résultat soit des connaissances que l'on a des processus de la production de la parole, soit de catégorisation phonémique des événements acoustiques. Le couple locuteur~auditeur joue également un rôle actif dans le traitement de la variation dans le signal acoustique.

Les études s'appuyant sur les « *Théories Actives* » de perception de la parole refusent une relation directe entre le signal acoustique et les unités linguistiques perçues. Liberman et ses collègues (voir Liberman, 1957 ; Liberman *et al.*, 1967_a ; Liberman *et al.*, 1967_b* ; etc...) ont mis en évidence quelques raisons expliquant le rejet d'un modèle passif :

- La durée moyenne d'un phonème affecte sa perception. Si elle est inférieure à un seuil précis (≈ 20 ms), il est impossible de reconnaître le phonème. Si l'on met une série d'événements acoustiques d'une durée équivalente à 20 ms les uns à la suite des autres et si on les fait écouter à un auditeur, celui-ci n'entendra qu'une sorte de bourdonnements à la place des phonèmes,
- Lors de la perception d'un phonème quelconque présenté dans différents contextes phonétiques, on perçoit toujours le même phonème. Autrement dit, pour des stimuli physiquement différents, on perçoit un seul stimulus correspondant à la catégorie phonémique de ce son,
- Si l'on procède à la segmentation de la parole, le résultat obtenu d'un segment quelconque ne correspondra pas à un phonème précis, car ce sont des chevauchements entre les différents segments qui constituent la parole :

« *The speech signal typically does not contain segments corresponding to the discrete and commutable phonemes [...] We cannot cut either the /di/ or the /du/ pattern in such a way as to obtain some pieces that will produce /d/ alone. If we cut progressively into the syllable from the right-hand end we hear /d/ plus a vowel or a nonsense sound; at no point will we hear only /d/.* »

(Liberman *et al.*, 1967_a : p. 436).

Toutes ces observations laissent supposer que la perception de la parole fait appel à un code spécial, à la place de l'utilisation de groupes de neurones représentant les phonèmes (comme présenté par certaines « *Théories Passives* » de perception de la parole). Liberman et ses collègues ont proposé que la perception de différents sons de la parole fasse intervenir un modèle articulatoire de production. Nous présentons dans la partie qui suit quelques « *Théories Actives* » qui mettent l'accent à la fois sur le rôle du couple locuteur~auditeur en perception, et sur les liens existant entre les mécanismes de production et de perception.

2.3.2.1. La Théorie Motrice de perception de la parole

Liberman *et al.* (1967_a, 1967_b*, 1972*) ; Liberman (1970) ; Liberman et Mattingly (1985) sont à l'origine de la *Théorie Motrice de perception* de la parole. Pour cette théorie, le locuteur~auditeur joue un rôle très important en production comme en perception. Pour les auteurs, la parole est perçue en référence à la production des phonèmes dans le signal acoustique. Ils postulent que nous percevons les sons en fonction des mêmes instructions neurologiques commandant nos articulateurs afin d'avoir une correspondance « parfaite » entre le signal acoustique produit et perçu.

Pour défendre leur théorie, les auteurs postulent, premièrement, que **la perception de la parole s'effectue au moyen d'une fonction spéciale**, car elle semble être liée à la structure grammaticale de la langue. Pour eux, le système auditif n'est pas préparé pour la

perception et le décodage de la parole, mais la relation qui existe entre la production et la perception des sons semble être en parfaite harmonie (Liberman *et al.*, 1972*)³².

Deuxièmement, ils postulent que **la perception de la parole est différente des autres mécanismes de perception sensorielle**. Les auteurs motivent cette considération en présentant le traitement et l'analyse des phonèmes comme dépendant d'une analyse en parallèle (comme en vision) et non en série. Autrement dit, le signal acoustique doit contenir beaucoup de petits morceaux d'information coexistant ensemble plutôt que des patterns discrets caractéristiques du phonème :

« The acoustic cues for successive phonemes are intermixed in the sound stream to such an extent that definable segments of sound do not correspond to segments at the phonemic level. Moreover the same phoneme is most commonly represented in different phonemic environments by sounds that are vastly different. »

(Liberman *et al.*, 1967_a : p. 432).

En troisième lieu, les auteurs postulent que, à l'inverse des « *Théories Passives* », pour lesquels le traitement de la parole s'opère en utilisant un système de gabarit de neurones représentant chacun un trait distinctif, **le décodage de la parole se fait d'une manière précise qui dépend de la façon dont nous produisons les sons de la parole**. Pour les auteurs, les traits distinctifs ne sont pas intégrés dans le signal acoustique, mais peuvent y être identifiés après reconstruction.

Quatrièmement, les auteurs de la *Théorie Motrice* postulent que **l'identification perceptive des phonèmes, qui est basée sur un pattern acoustique précis, serait conduite par l'auditeur**. D'après eux :

³² Mattingly (1972*) propose que la langue soit composée de deux systèmes : les représentations sémantiques générées par l'intellect et les « *Social Releasers* » comprises dans l'articulation des sons. L'auteur explique que ce sont les impédances entre les deux systèmes qui ont nécessité l'évolution de la grammaire de la parole. Le code de la parole, avec l'aide de différentes conversions complexes, comble parfaitement son traitement sur les deux axes : sémantiques et acoustiques.

« We shall assume – indeed we think we must assume – that somewhere in the speaker's central nervous system there exist signals which stand in a one-to-one relation to the phonemes of the languages. In the act of speaking, these signals, arranged of course in some temporal pattern, flow outward from the central nervous system and eventuate as commands to the articulatory muscles. »

(Liberman *et al.*, 1967^{b*} : p. 77).

Pour les auteurs, l'invariant n'existe pas en tant que tel dans le signal acoustique, mais il est dérivé des connaissances que l'on a du système de production de la parole. Il est donc à noter que la parole est encodée suivant des fonctions spécifiques qui trouvent leurs bases dans les mécanismes articulatoires. Les auteurs considèrent qu'il faut traiter le signal acoustique à l'envers afin de retrouver les informations concernant l'encodage de la parole. Ils présentent la clé nécessaire à décoder le signal acoustique en se basant sur une représentation neurologique des traits, qui comprend également l'activité de chevauchement de différents réseaux de neurones (voir Figure 16). À partir de ce schéma de production, la perception de la parole peut être dérivée.

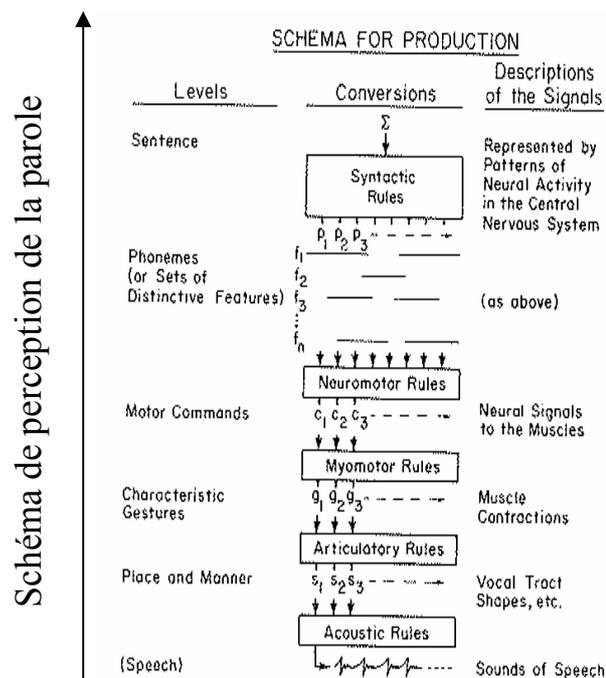


Figure 16 : Représentation schématique des étapes proposées en production de la parole.
(Adaptée de Liberman *et al.*, 1967^a).

Et enfin, les auteurs postulent que, comme pour le *Modèle Neurologique* (voir les « *Théories Passives* »), **la perception fonctionne avec des éléments spécialisés dans le**

décodage de la parole. Néanmoins, ils ne sont pas d'accord avec les partisans des « *Théories Passives* » sur l'identité des éléments traités par les mécanismes physiologiques spécialisés. Pour Liberman *et al.* (1972*), ce ne sont pas des cellules neurologiques spécifiques qui permettent le décodage de la parole, mais d'autres éléments :

« *In sum, there is now a great deal of evidence to support the assertion that man has ready access to physiological devices that are specialized for the purpose of decoding the speech signal and recovering the phonetic message. Those devices make it possible for the human being to deal with the speech code easily and without conscious awareness of the process of its complexity.* »

(Liberman *et al.*, 1972*: p. 322).

Ce mode spécifique de perception de la parole défendu par les partisans de la *Théorie Motrice de la perception* a été démontré dans des expériences sur la perception duplex « *Duplex perception* ». Liberman *et al.* (1981) ont montré que les auditeurs perçoivent deux stimuli simultanément, dans une expérience où la transition du F₃ a été isolée et présentée à une seule oreille et le reste de la syllabe à l'oreille opposée (voir Figure 17). Les auteurs ont rapporté que si la syllabe synthétisée en Figure 17, partie C, est présentée sans la transition de F₃, (voir Figure 17, partie B), les auditeurs perçoivent une syllabe synthétique CV avec un lieu d'articulation indéterminé. Ceci indique que comme différentes informations acoustiques sont nécessaires à l'identification du lieu d'articulation d'une consonne, on s'attendrait à observer les mêmes incidences sur l'identification des voyelles. En effet, et comme nous l'avons vu dans les parties précédentes, les approches « *Elaborated Target* » et « *Dynamic Specification* » postulent que ce sont différentes informations présentes dans le signal qui déterminent la « couleur » de la voyelle et permettent ainsi son identification. Pour percevoir une consonne avec un lieu d'articulation précis, on a besoin de l'information contenue dans la transition de F₃ en plus de celle de F₁ et F₂ pour permettre son identification. Pour identifier les voyelles coarticulées avec les consonnes, on a besoin de l'information contenue dans les transitions de F₁, F₂ & F₃ pour permettre son identification, car les transitions font partie intégrante des voyelles.

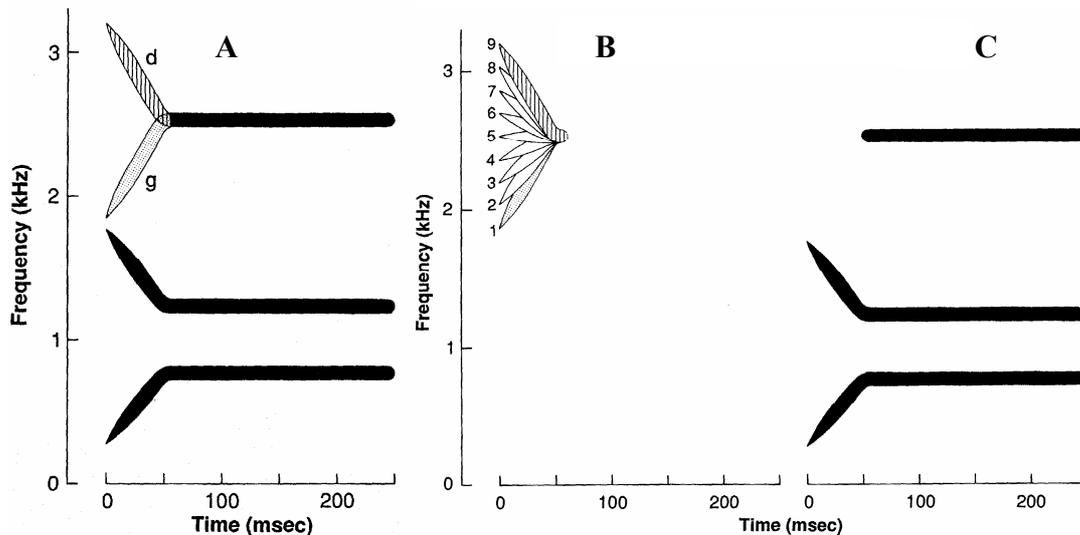


Figure 17 : Schémas représentant la production d'une syllabe CV originale pour distinguer entre un [da] et [ga] (en A) et d'une syllabe CV synthétisée, utilisée dans l'expérience de perception duplexe, divisée en deux parties : une série de transitions de F₃ couvrant les plages allant de [da] et [ga] (en B) et le reste de la syllabe (en C). (Emprunté à Liberman *et al.*, 1981).

Pour cette « *Théorie Active* », l'unité primitive en perception de la parole est le *Geste Articulatoire* (i.e., des *Syllable-Sized units*), qui naît du chevauchement de différents traits constituant les phonèmes³³. L'invariant n'existe pas dans le signal acoustique, mais il est présent dans le signal acoustique à partir des traits constituant les *Syllable-Sized units* : ce sont des informations **dynamiques** contenues dans le signal acoustique qui permettent ainsi de l'identifier.

Un autre modèle faisant partie des « *Théories Actives* » est le *modèle de l'analyse par synthèse* proposé par Stevens et collègues. Comme on le verra dans la partie qui suit, ce modèle ne partage pas toutes les hypothèses de la *Théorie Motrice*.

2.3.2.2. Le modèle de l'Analyse par Synthèse

Stevens (1960), Stevens & Halle (1967*) et Stevens & House (1972*) ont développé ce modèle de perception de la parole, qui comme le suppose la *Théorie Motrice*, présente le

³³ Les auteurs expliquent que lorsqu'on produit le mot [bæg], les $\frac{2}{3}$ de la voyelle [æ] sont modifiés par la coarticulation des consonnes [b] et [g], considérant ainsi que l'unité primitive de la perception de la parole doit être considérée comme ayant une partie initiale (l'**onset** consonantique), une partie médiane (la partie « stable » de la voyelle) et une partie finale (l'**offset** consonantique).

couple locuteur~auditeur comme ayant un rôle très important dans les deux mécanismes de production et de perception de la parole. Pour eux, les différentes informations dérivées du signal acoustique sont structurées par le processus articulatoire. Les auteurs postulent que :

« [...]the perception of speech involves the internal synthesis of patterns according to certain rules, and a matching of these internally generated patterns against the pattern under analysis. We suggested moreover, that, the generative rules utilized in the perception of speech were in large measure identical to those utilized in speech production, and that fundamental to both processes was an abstract representation of the speech event. »

(Stevens & Halle, 1967 : p. 88).

Les auteurs pensent que le signal acoustique est décodé en unités ayant des représentations abstraites des segments et des traits utilisant les mêmes règles phonologiques que celles que l'auditeur emploie dans la production du signal acoustique. En effet, ils donnent un poids important au système auditif et aux connaissances phonologiques de la langue qui facilitent le décodage du signal par l'auditeur. Celui-ci utilise ses compétences phonologiques afin de générer une image mentale synthétisée du son. Il tente ensuite de faire correspondre l'input acoustique avec l'hypothèse mentale synthétisée qu'il a des sons. La décision phonétique que l'auditeur prend en compte, est établie sur la base d'une comparaison entre trois types d'information fournis par ce modèle (voir Figure 18).

Suivant les propositions de la *Théorie Motrice*, pour les partisans de la *Théorie de l'Analyse par Synthèse*, les informations concernant le mécanisme de production du signal acoustique peuvent être utilisées par l'auditeur dans le mécanisme de décodage de la parole.

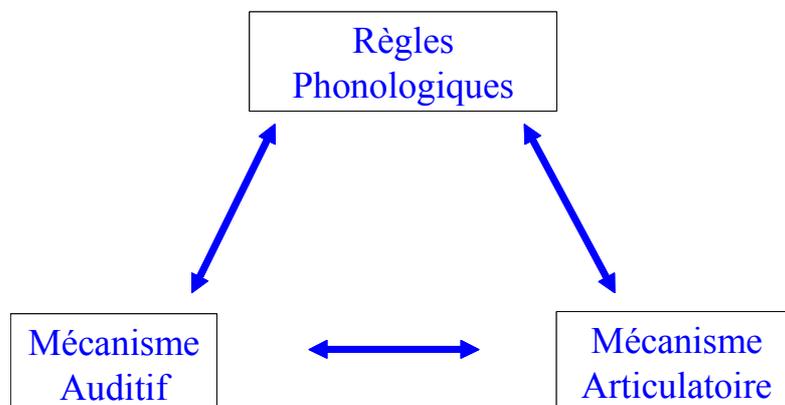


Figure 18 : Modèle du mécanisme articulatoire et du processus de génération de la parole.
(Inspiré de Stevens & Halle, 1967*).

Pour cette théorie, l'invariance peuvent être présentes dans le signal acoustique pour quelques types de sons (e.g., l'invariance du « locus virtuel » ou de l'**équation de locus** quant à la détermination du lieu d'articulation des consonnes adjacentes aux voyelles, voir § 4.3.2.1.1). Pour cette théorie, l'unité primitive de perception de la parole est le trait distinctif.

2.3.2.3. La Théorie Quantique

Kenneth Stevens a proposé la *Théorie Quantique* (voir Stevens, 1972, 1989, 2003). Cette théorie n'est pas un modèle de perception de la parole en tant que tel, et n'est ni un modèle actif ni passif, car aucune référence explicite du rôle actif ou passif de l'auditeur dans le traitement de la perception de la parole n'est donnée. Néanmoins, les suppositions que fait Stevens sont d'une grande importance pour la compréhension du mécanisme de perception de la parole.

En se basant sur les propositions de la *Théorie Acoustique de la Parole* (Fant, 1960), Stevens explique que, pour un paramètre articulaire précis, des changements articulaires mineurs produits dans certaines régions conduiront à des changements acoustiques importants tandis que dans d'autres régions, des changements articulaires majeurs produiront des changements acoustiques mineurs. Pour cette théorie, le rôle attribué à l'auditeur est tout simplement d'arriver à dériver les traits constituant le signal acoustique présentent dans des régions précises dans le conduit vocal (voir Figure 19) et les mouvements de différents paramètres acoustiques constituant ces traits. Une relation « *one-to-one* » est alors automatiquement établie entre les différents événements de la production de la parole et les unités de perception, d'une manière entièrement passive.

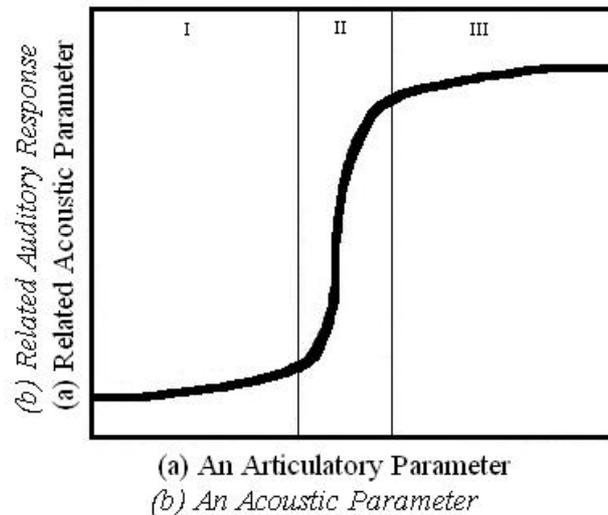


Figure 19 : Courbe représentant schématiquement les principes de base de la Théorie Quantique. Les changements causés par les paramètres acoustiques, dus aux changements réguliers des paramètres articulatoires sont présentés sur les deux axes en (a). Les axes en italiques (sous (b)), représentent les réponses auditives dues aux changements réguliers des paramètres acoustiques. (Emprunté à Stevens, 1989).

La forme sigmoïdale de la courbe reflète la relation non-linéaire entre les mouvements articulatoires et leurs corrélats acoustiques. Les différences entre les trois zones s'expliquent par le fait que dans la région I, les changements des paramètres articulatoires produiront des changements acoustiques négligeables. Dans la région II, des changements importants des paramètres articulatoires produiront des modifications importantes de l'output acoustique. Et enfin, dans la région III, l'output acoustique est de nouveau insensible aux changements articulatoires, bien que la réponse acoustique soit différente de celle de la région I :

« [...] The difference in the value of acoustic parameter from region I to region III is large. That is, there is a significant acoustic contrast between these two regions, which are separated by the intermediate region II in which there is a rather abrupt change in the acoustic parameter. [...] the difference in the acoustic pattern between regions I and III should not be regarded as simply a matter of identifying two points on a scale of some acoustic parameter [...] »

(Stevens, 1989 : p. 4).

Les régions I et III sont des régions de plateaux acoustiquement stables. La région II en revanche est une région de transition abrupte du paramètre acoustique traduisant une zone d'instabilité et de contraste avec les régions I et III.

Pour ce qui est de la relation entre les variations acoustiques et les réponses auditives de l'output, on suit le même schéma qu'en production. Autrement dit, les régions I et III sont

des régions auditivement stables et la région II est une région de transition du paramètre auditif et de contraste (i.e., une instabilité). D'après Stevens, la région II affecte largement la perception car les sons qui y sont produits nécessitent une grande précision articulatoire :

« [...] Region II can, in some sense, be considered as a threshold region such that as the acoustic parameter changes through this region the auditory response shifts from one type to another. »

(Stevens, 1989 : p. 4).

Pour Stevens, la région II est donc une région de discrimination et de contraste de sons (région où le seuil d'identification des sons se situe) : le changement rapide des caractéristiques du son au niveau acoustique provoque chez les auditeurs un changement de réponse aux stimuli (car situé dans le seuil d'identification) puisque ceux-ci sont perçus comme différents. L'auteur estime que dans cette région, la distribution des sons à l'intérieur d'une langue et leurs structures phonologiques sont expliquées, car il est possible de la rattacher à certains mécanismes du système auditif.

Du point de vue acoustique, les deux régions I et III, dites de stabilité, sont considérées comme les deux régions idéales pour la production des voyelles. D'après Stevens, les trois voyelles extrêmes /i a u/ ou « *Hot-Spots* » sont produites dans les zones de stabilité du conduit vocal. Elles sont considérées par Stevens comme des voyelles acoustiquement stables, ayant des formes perceptivement discriminables entre elles en comparaison aux autres voyelles situées dans la région d'instabilité articulatoire dans la région II. Cette stabilité des trois voyelles extrêmes /i a u/ repose sur une propriété quantique de ces voyelles : la prééminence dans le spectre vocalique explicitée par la proximité de deux ou plusieurs formants vocaliques, indiquant une production articulatoirement stable des voyelles (par exemple le rapprochement entre F_3 et F_4 pour /i/). La stabilisation de la voyelle /u/ se fait par le biais de l'arrondissement des lèvres et le rapprochement entre F_1 et F_2 et celle de la voyelle /a/ par la valeur élevée de son premier formant (se rapprochant de la limite extrême de la fréquence du premier résonateur). Nous reviendrons ultérieurement sur cette question.

Pour ce modèle, l'invariance est répartie en fonction de différents paramètres, tels que : le mécanisme articulatoire, le signal acoustique, le mécanisme de perception de la parole.

Ce modèle apporte quelques explications quant à la relation entre les mécanismes articulatoires et acoustiques d'un côté et acoustiques et auditifs de l'autre. Néanmoins, il n'est pas toujours facile de mettre en évidence cette relation au niveau de la stabilité et du contraste dans les trois régions définies dans la Figure 19 (pour les critiques du modèle, voir Stevens, 1989 ; Vallée, 1994).

2.3.2.4. La Théorie H&H

La *Théorie H&H* (pour *Hypo-* et *Hyper-Articulation*) a été proposée par Lindblom (1990) pour expliquer la relation entre les mécanismes de production et de perception de la parole. Cette théorie est considérée comme un modèle actif mettant en évidence le rôle du couple locuteur~auditeur en production et perception de la parole.

Le premier postulat de cette théorie est l'absence d'invariance articulatoire ou acoustique. Au niveau de l'invariant articulatoire, Lindblom montre, sur la base d'expériences de « *bite-block* » (dans ce type d'expérience, le locuteur bloque l'ouverture de son conduit vocal par un stylo ou une pipe placé entre ses dents), que la parole produite par ces locuteurs reste intelligible, malgré le blocage des articulateurs qui conduit à une production des sons sans atteindre la cible articulatoire normale. On observe en effet, un ajustement et une compensation du blocage du conduit vocal pour les autres articulateurs qui semble être un phénomène normal pour faciliter la compréhension du message.

« [...] Good performance on compensatory tasks is a consequence of the fact that motor behaviour in general is organized to be inherently compensatory ».

(Lindblom, 1990 : p. 419).

Ceci montre que certains sons peuvent être produits avec des configurations articulatoires différentes du conduit vocal, ayant comme conséquence la non-invariance de ces articulations.

Lindblom va également à l'encontre de l'existence de l'invariant acoustique. En s'appuyant sur les résultats qu'il a obtenus sur la réduction vocalique « *Vowel Undershoot* » et sur le phénomène de coarticulation. Pour lui, les voyelles ont des propriétés acoustiques

différentes en fonction du contexte environnant, mais sont parfaitement intelligibles dans les différents débits et styles de parole grâce à leurs entourages consonantiques et/ou syllabiques.

La deuxième supposition que fait Lindblom dans sa théorie concerne le mécanisme de perception de la parole. En effet, partant du premier postulat que les invariants acoustiques et articulatoires sont absents, la perception de la parole semble dépendre de différents facteurs autres que ceux inclus dans le signal acoustique. Pour Lindblom, il existe donc deux sources d'information que l'auditeur utilise dans le processus de perception de la parole : i) l'information contenue dans le signal et ii) l'information externe au signal.

Cette théorie est basée sur deux principes : la **plasticité** et l'**économie** que le locuteur et l'auditeur utilisent en parallèle. Lindblom explique la **plasticité** en parlant de l'adaptation mutuelle qu'ont les locuteurs et les auditeurs. Le locuteur peut changer sa façon de parler en suivant un continuum allant de l'*Hypo-Articulation* à l'*Hyper-Articulation* afin de s'adapter à la situation communicationnelle (voir Figure 20). L'auditeur suit également le même principe en attribuant plus ou moins d'importance aux informations véhiculées par le signal acoustique. Pour ce qui concerne l'**économie**, Lindblom explique qu'en général, le mécanisme de production de la parole tend à réduire les efforts (i.e., se rapprochant d'une forme *Hypo-Articulée*), à moins que la situation contextuelle implique une attention plus importante à la qualité du signal acoustique pour permettre une meilleure compréhension du message.

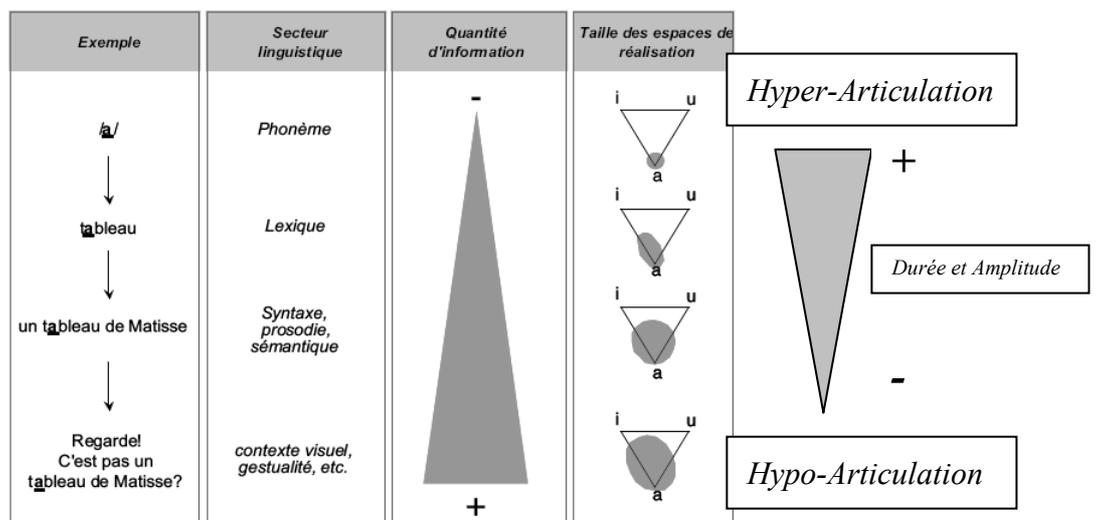


Figure 20 : Variations phonétiques en fonction de la quantité d'information contextuelle véhiculée. (Inspiré de Meunier, 2005 ; adapté de Lindblom, 1990).

D'après l'auteur, le locuteur se base sur ses connaissances des formes lexicales, grammaticales, et linguistiques, ainsi que sur les caractéristiques contextuelles de sa langue lors de la communication afin de transmettre un message compréhensible à l'auditeur. Un mot ou une phrase produite par un locuteur peut ne pas contenir toutes les informations nécessaires à sa compréhension, mais le peu d'information que le message contient peut suffire à comprendre les intentions de celui-ci. Ainsi, plus la quantité d'information utilisée (venant des niveaux morphologique, syntaxique et sémantique) est importante, plus la précision articulatoire lors de la production d'un segment est basse et par conséquent, plus la variabilité contextuelle est importante. La compréhension d'un message se base donc sur les connaissances *a priori* qu'un locuteur a de sa langue. L'articulation se situerait ainsi sur un continuum allant de l'*Hyper-Articulation* à l'*Hypo-Articulation* (voir Figure 20). Ainsi, plus l'amplitude et la durée des gestes articulatoires sont grandes, plus la production du locuteur est proche d'une forme *Hyper-Articulée*. En revanche, plus les informations sur le contexte seront nombreuses et moins le locuteur devra maintenir une clarté élevée lors de l'articulation. Le style *Hypo-Articulé* naît de la coarticulation de différents segments de la chaîne parlée : la coarticulation donne lieu à des phénomènes de réduction vocalique (i.e., « *Undershoot* » vocalique, Lindblom, 1963_a) lesquels se traduiront par une relative imprécision des gestes articulatoires.

Le rôle du locuteur dans la transmission du message est important car il adapte sa production en fonction de la compréhension de l'auditeur. La production d'une forme isolée d'une voyelle conduirait à une variabilité moindre au niveau intra-individuel (mais plus ou moins grande au niveau inter-individuel du fait des variations de taille de conduit vocal) et la forme *Hyper-Articulée* de cette voyelle se rapprocherait davantage d'un prototype qui correspond plus ou moins à une production vocalique de type cardinal. En revanche, la production de la même voyelle dans un environnement consonantique produit une variabilité importante du fait de la coarticulation de différents segments, et s'éloigne donc du prototype vocalique.

L'auditeur va chercher à comprendre le message véhiculé par le locuteur en normalisant les variations. Généralement, l'interprétation du message se déroule sans

problème, mais parfois des erreurs d'interprétation conduisent à une mauvaise compréhension du message et produisent des changements linguistiques dans le temps.

Pour Lindblom, le locuteur et l'auditeur sont en parfaite harmonie, ils s'adaptent mutuellement aux situations communicationnelles diverses. Cette adaptation se situe dans un continuum allant d'une forme *Hypo-Articulée* à une forme *Hyper-Articulée*. C'est donc cette adaptation qui va expliquer en partie les deux mécanismes de production et de perception de la parole.

En effet, le locuteur, lors de sa production, va produire beaucoup de chevauchements entre les différents segments constituant le message que ce qui est acceptable pour l'auditeur. Ce dernier va tenter de comprendre le message dans le contexte de la situation communicative. Les connaissances et les attentes que l'auditeur a de ce qui doit être dit par le locuteur ont la même importance que le message lui-même. Les mots produits par le locuteur apportent tout simplement l'information qui manque à la situation communicationnelle, interagissant avec la compréhension par l'auditeur du propos général du message et de l'attitude émotionnelle véhiculé par le locuteur. Lindblom appelle cette situation : l'approche « *signal-plus-knowledge* »³⁴ :

« The way one views the task of the speaker depends on one's assumptions about how speech perception works. Consequently, we shall begin by considering the following picture of perceptual processes. The first observation is that the perception of speech is set up to cope with partial information. That seems to be a general design feature of perceptual systems ~Shepard, (1984)!. The reason why speech perception succeeds in coping with partial information is that percepts are never completely "raw" records of the physical signal. They are inevitable products of signal-plus-knowledge. This is true not only of speech but of perception in general. Also it is true not only for "higher" levels of language processing, but even for the most elementary tasks of speech perception, e.g., recognition of nonsense CV syllables... Second, speech perception does not presuppose the presence of invariants in the acoustic signal[...] ».

(Lindblom, 1996 : p. 1684).

³⁴ Pour Lindblom, le *plus* dans l'approche "*signal-plus-knowledge*" signifie interagissant avec (voir Lindblom, 1996, note : 6, pp. 1691).

La conséquence directe des arguments apportés par Lindblom, par rapport aux interactions entre locuteur et auditeur, est l'abandon de la recherche d'une simple unité primitive de perception de la parole. La raison que Lindblom avance est qu'aucun niveau (articulatoire, acoustique ou perceptif) n'est privilégié. Mais il avance cependant l'idée selon laquelle le niveau perceptif peut avoir la priorité car il porte toute l'information sensorielle cruciale à l'identification de la parole.

Nous avons vu que pour Lindblom, le locuteur adapte sa production en fonction de la pertinence de l'information véhiculée. Sa production se situera dans un continuum allant de l'*Hypo* à l'*Hyper-Articulation*. Mais c'est davantage la forme *Hypo-Articulée* qui est privilégiée. Cette forme correspond à une réduction articulatoire où une cible n'est pas atteinte. Généralement, cette réduction articulatoire n'est pas perceptible : l'interaction entre les gestes et la durée très courte des segments conduit à l'absence de gestes au niveau acoustique. Au niveau articulatoire, les résultats obtenus par Browman & Goldstein (1990, 1992) dans leur modèle de *Phonologie Articulatoire*, montrent que ces cibles non-atteintes au niveau acoustique semblent être réalisées par les locuteurs au niveau articulatoire. C'est le sujet de la partie qui suit.

2.3.2.5. Le rôle des « hidden gestures » en Phonologie Articulatoire

En *Phonologie Articulatoire*, les gestes articulatoires sont considérés comme des événements physiques, discrets, qui se déroulent dans le processus de production de la parole. La *Phonologie Articulatoire* tend à décrire les unités lexicales en fonction de ces événements et de leur interrelation en production de la parole. Ces unités lexicales qui constituent la parole sont représentées par des gestes articulatoires coordonnés (ou des constellations de geste). Chacun de ces gestes articulatoires correspondent à des déformations du conduit vocal considérée(s) comme linguistiquement significative(s), par exemple, la fermeture ou la protrusion labiale, l'avancement du dos de la langue, l'abaissement vélaire, etc...

Browman & Goldstein (1990, 1992) ont montré qu'en style « *casual speech* » (i.e., style relâché à débit rapide), tous les gestes articulatoires sont réalisés avec succès. D'après eux, la réalisation des gestes articulatoires se fait en augmentant le chevauchement temporel entre ces gestes et en réduisant leur ampleur temporellement et spatialement. Les auteurs ont travaillé sur des phénomènes d'assimilation consonantique en utilisant la radiographie X. Ils donnent comme exemple la réalisation du [t] final du mot « *perfect* » lors de la production de l'expression « *perfect memory* » à un débit normal et à un débit rapide. Les résultats radiographiques du mouvement vertical du dos de la langue, de l'apex de la langue et de la lèvre inférieure, montrent qu'en débit normal (voir Figure 21, partie a)), la production du [t] final du mot « *perfect* » est complètement audible et visible sur le signal acoustique. On observe la réalisation des gestes vélaire (κ) pour [k] et alvéolaire (τ) pour [t] presque simultanément, indiquant que la fermeture alvéolaire n'est réalisée qu'une fois la fermeture vélaire produite. Il est à noter que l'onset du geste alvéolaire est caché (« *hidden* »), bien que sa réalisation soit audible. Le geste labial (β) pour [m] dans le mot « *memory* » quant à lui est réalisé quelques millisecondes après le relâchement du geste alvéolaire.

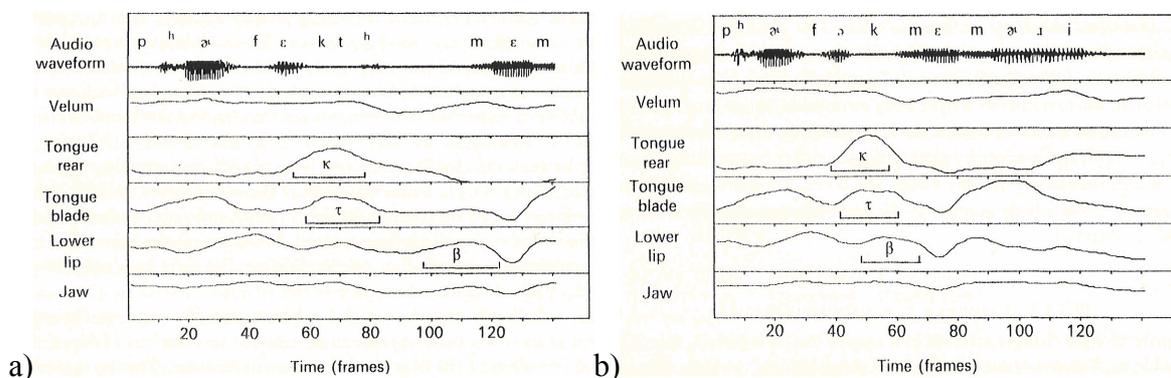


Figure 21 : Trajectoires de « pellets » en radiographie X de l'expression « *perfect memory* » réalisée en débit normal (en a) et en débit rapide (en b)). κ = geste de fermeture vélaire, τ = geste de fermeture alvéolaire et β = geste de fermeture labiale. (Emprunté à Browman & Goldstein, 1990).

En ce qui concerne la réalisation en débit rapide (voir Figure 21, partie b)), le [t] final de « *perfect* » a subi une élision complète, aucune impression auditive et aucun relâchement acoustique ne sont observés. Mais en observant les radiographies X de la réalisation articulatoire de « *perfect memory* », on remarque l'existence du geste alvéolaire (τ) pour [t], ayant quasiment la même magnitude articulatoire qu'en débit normal, qui est réalisé simultanément avec le geste vélaire (κ) pour [k]. La compression temporelle affecte la

réalisation du geste labial (β) pour [m] dans le mot « *memory* ». On observe en effet une réalisation quasi simultanée des gestes articulatoires labial (β) et vélaire (κ). Le geste alvéolaire (la fermeture et le relâchement) subit un chevauchement lors de sa réalisation avec les gestes vélaire et labial, mais aucune présence acoustique n'est démontrée.

Pour les auteurs, les gestes articulatoires ne subissent aucune élision spatiale, mais leur correspondance acoustique n'est pas audible. Une réduction vocalique comme décrite par Lindblom (1963_a) peut être alors obtenue par compression temporelle, bien que tous les gestes articulatoires réduits soient réalisés.

Si ce genre de « cache perceptuel » peut être démontré systématiquement, cela pourrait soutenir les théories pour lesquelles la perception utilise les gestes comme unité primitive de base, bien qu'ils soient non perceptibles par les auditeurs. Dans leur travail, Browman & Goldstein (1990, 1992) mettent l'accent sur l'analyse de la forme *Hypo-Articulée* des unités lexicales, l'étude de la forme *Hyper-Articulée* de ces unités lexicales n'est pas démontrée. Pour eux en effet, les processus d'*Hyper-Articulation* peuvent être obtenus dans leur modèle par chevauchement décroissant et par augmentation de l'ampleur des gestes.

La plupart des études développées dans cette partie, traitant des interactions entre les mécanismes de production et de perception de la parole, visent à identifier les unités perceptives et leurs caractérisations. Elles ne se sont pas penchées sur les représentations perceptives des auditeurs. Dans les parties qui suivent, nous nous intéressons au modèle de Patricia Kuhl et ses collègues (*The Native Language Magnet Model*) puis à la perception des prototypes proposée par Kenneth Johnson.

2.3.2.6. Le Native Language Magnet model

Patricia Kuhl et ses collègues ont montré que les adultes étaient capables d'identifier le **prototype**³⁵ phonétique de la voyelle /i/ en anglais américain, i.e., ils arrivent à identifier la meilleure instance de la catégorie phonémique de /i/. Les auteurs montrent que les adultes et

³⁵ Pour Rosch (1975*), le **prototype** est une abstraction issue de la distribution statistique des instances d'une catégorie. Il est considéré comme la meilleure instance d'une catégorie.

les bébés de 6 mois ne sont pas capables de discriminer correctement les voyelles auditivement très proches de celle de la voyelle **prototypique** /i/ choisie par les adultes. Ces résultats ont été obtenus dans une expérience menée par Kuhl & Iverson (1995*), où des adultes et des bébés de 6 mois ont eu comme tâche la discrimination (décision identique ou différente) entre des voyelles synthétisées regroupées par paire. Deux listes ont été présentées aux auditeurs. Dans la première, la comparaison s'effectuait avec une voyelle /i/ **prototypique** et dans la seconde la voyelle /i/ était **non-prototypique**³⁶. Les auteurs montrent que le taux de discrimination entre les paires vocaliques est significativement meilleur pour la liste contenant des voyelles **non-prototypiques** (Kuhl & Iverson, 1995*). Kuhl montre que cet effet de baisse de discrimination entre les voyelles proches du prototype peut s'expliquer en se basant sur « *The Perceptual Magnet Effect* » (Kuhl, 1991). Le **prototype** fonctionne comme un aimant pour les autres stimuli appartenant à la même catégorie phonémique (la catégorie de /i/, ici). Lorsque les membres d'une catégorie se trouvent à proximité d'un **prototype** (voir Figure 22, A), ceux-ci sont attirés vers celui-ci (voir Figure 22, B).

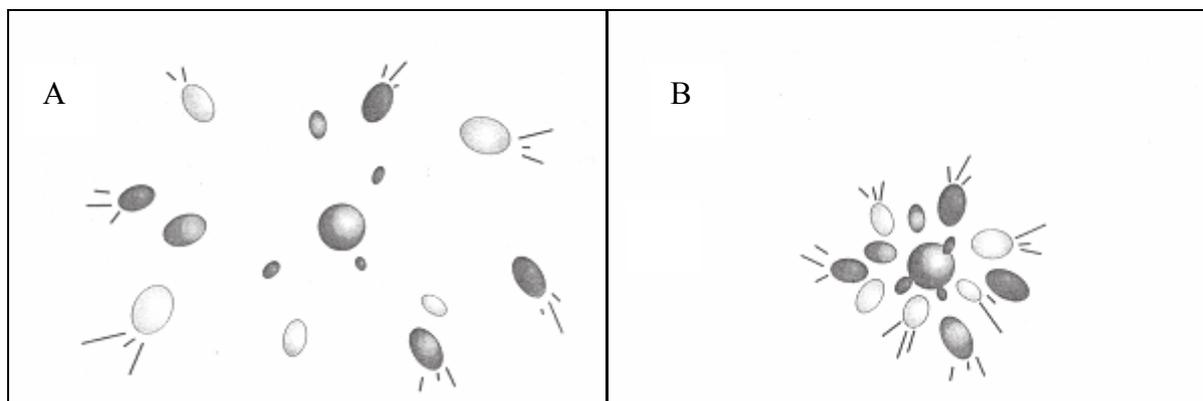


Figure 22 : Schéma montrant l'effet magnétique perceptif du prototype. Les membres de la catégorie se trouvent à proximité du prototype (en A), et le prototype les attire vers lui (en B). (Emprunté à Kuhl, 1994).

Pour Patricia Kuhl, « *The Perceptual Magnet Effect* » va structurer l'organisation des catégories phonémiques à l'intérieur de l'espace vocalique. Iverson et Kuhl (1995) ont montré que « *The Perceptual Magnet Effect* » est :

³⁶ Il est à noter que toutes les voyelles présentées aux auditeurs possèdent des caractéristiques très proches de la voyelle /i/ et les distances acoustico-auditives entre les paires de voyelles étaient toujours les mêmes dans les deux listes.

- Présent chez les adultes comme chez les bébés de 6 mois,
- Sensible aux expériences linguistiques (des enfants américains et suédois montrent un effet magnétique de leurs voyelles uniquement),
- Non-présent chez les singes, et
- Associé à la réduction de la discriminabilité autour du prototype.

Les conclusions des auteurs montrent que l'acquisition d'une langue requiert une déformation de l'espace perceptif des catégories phonétiques, conduisant à une réduction de la distance perceptive autour du **prototype**. En effet, Patricia Kuhl a proposé dans son modèle : « *The Native Language Magnet* » que l'espace acoustique est divisé en frontières psycho-acoustiques correspondant aux caractéristiques phonémiques de chaque langue, qu'elle appelle la carte perceptive (voir Figure 23). Hawkins (1999_c) explique à quoi correspondent ces frontières psycho-acoustiques (voir Kuhl & Iverson (1995*)) pour plus de détails sur ce modèle) :

« [...] Kuhl proposes that the baby is innately endowed with the ability to divide the speech signal into gross categories, the "basic cuts" of the speech sound space. These basic categories are typically more detailed than those used by any single language, because they encompass all gross contrasts that languages could potentially employ. Because animals can also distinguish some of the categories that babies can, Kuhl suggests that the boundaries of these gross categories are natural psychoacoustic boundaries ».

(Hawkins, 1999_c : p. 250).

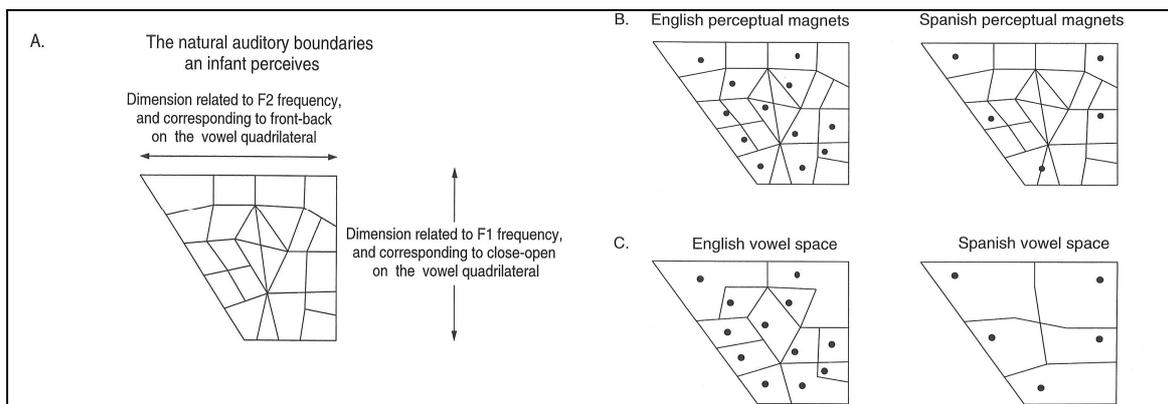


Figure 23 : Schéma montrant le développement de l'espace acoustique des voyelles des bébés en suivant le modèle de « Native Language Magnet » proposé par Patricia Kuhl. (Emprunté à Iverson & Kuhl, 1995).

Kuhl propose que, à la naissance, les bébés aient la capacité innée de percevoir toutes les frontières naturelles qui sous-tendent les différences entre les catégories vocaliques, à l'intérieur desquelles la distribution des timbres vocaliques n'est pas construite. Ceci indique que « *The Perceptual Magnet Effect* » des meilleures instances des voyelles n'est pas encore développé. La partie A de la Figure 23 montre les frontières auditives naturelles de l'espace acoustique vocalique proposées dans ce modèle. À 6 mois, « *The Perceptual Magnet Effect* » spécifique à chaque langue est construit. Le bébé est capable de discriminer toutes les instances des catégories vocaliques de sa langue maternelle. Il reste néanmoins sensible aux différentes catégories non utilisées dans sa langue maternelle. La partie B de la Figure 23 montre l'exemple de « *The Perceptual Magnet Effect* » des **prototypes** vocaliques en anglais et en espagnol. À 10-12 mois, le bébé n'est plus sensible aux différentes catégories ne faisant pas partie de son système vocalique. La partie C de la Figure 23 montre les changements des frontières des catégories dans les deux langues, l'anglais et l'espagnol. Les divisions des espaces acoustiques correspondent à celles des catégories phonémiques.

Patricia Kuhl montre que la première représentation de la parole, chez les enfants, est entièrement auditive, i.e., passive. Néanmoins, cette représentation englobe différents éléments visuels, moteurs, mais également des correspondances établies à partir des sons produits par les adultes qui l'entourent, ainsi que par les sons qu'il peut produire lui-même avec les contraintes articulatoires de son conduit vocal. La perception de la parole est alors considérée comme multimodale et influencée par l'expérience que le bébé acquiert au fur et à mesure de l'apprentissage des catégories phonémiques de sa langue maternelle. Et comme la perception de la parole est organisée en suivant le développement des structures spécifiques de sa langue, la perception influence petit à petit la production de la parole chez l'enfant.

Ce modèle décrit les interactions entre les capacités innées du système auditif et l'apprentissage des catégories phonémiques spécifiques à une langue, par le biais de l'étude des catégories phonétiques produisant les changements dans les représentations mentales des catégories chez les bébés.

Le modèle proposé par Patricia Kuhl n'explique pas directement la nature des **prototypes** phonétiques, l'effet contextuel sur les **prototypes** ou encore la réorganisation perceptive chez les enfants. Les auteurs expliquent que la réorganisation perceptive des représentations mentales chez le bébé peut intervenir soit à un stade tardif de la perception de la parole (i.e., après traitement des catégories) soit au cours du développement de son système auditif. Patricia Kuhl utilise le mot **prototype** d'un point de vue neutre, i.e., sans prendre partie ni pour une série d'éléments constituant la représentation du phonème dans le cerveau, ni pour une unité abstraite de la représentation mentale du phonème. Elle explique qu'il est possible de fusionner les deux en ayant une représentation abstraite du phonème et une série d'éléments desquels l'unité abstraite peut être dérivée. Cette vision montre que les deux mécanismes de production et de perception de la parole peuvent évoluer dans le temps, chez les bébés et chez les adultes. Ainsi, par exemple, un adulte est capable d'adapter facilement sa production et sa perception à un nouvel accent régional.

Dans leurs expériences perceptives pour identifier les **prototypes**, Kuhl & Iverson (1995*) ont utilisé des voyelles produites en isolation. Or on sait que les propriétés perceptives des voyelles produites en isolation sont très différentes de celles produites en contexte. À ce propos, différentes études ont mis en évidence le rôle de « *The Perceptual Magnet Effect* » et les unités **prototypiques** perceptives des allophones.

Barrett (1998*) a montré que « *The Perceptual Magnet Effect* » existe en perception des allophones de la voyelle /u/ : /u/ en isolation, et en contexte /lu/ et /ju/, en anglais britannique. En production de la parole, la réalisation de la voyelle /u/ est différente dans les trois structures : une augmentation de la valeur de F_2 allant de /u/ en isolation jusqu'à /ju/ en passant par la réalisation de /lu/. En test de discrimination, Barret (1998*) a montré qu'il existe trois « *Perceptual Magnet Effects* » distincts pour chaque allophone, i.e., chaque valeur de F_2 n'est utilisable que pour un contexte précis. L'auteur explique que les voyelles synthétiques représentant ces trois allophones doivent avoir des fréquences de F_2 différentes en fonction du contexte consonantique. Ceci montre alors que les auditeurs ont des « représentations » différentes des voyelles coarticulées par rapport aux voyelles produites en isolation.

Les résultats obtenus par Barret (1998*) montrent des implications intrigantes (Hawkins, 1999_c). En effet, si « *The Perceptual Magnet Effect* » implique l'existence d'un **prototype** phonétique qui joue un rôle important dans le mécanisme de perception de la parole, on doit être confronté à l'un des deux cas de figures suivants. Le premier cas de figure indique l'existence d'un nombre illimité de **prototypes** phonétiques sensibles aux variations contextuelles stockés dans le cerveau (voir également Wickelgren, 1969* ; qui peut provoquer un grand chevauchement perceptif au niveau des **prototypes**, ainsi qu'un affaiblissement des liens existant entre les **prototypes** d'un côté et « *The Perceptual Magnet Effects* » de l'autre. Le second cas de figure en revanche, pour qui les **prototypes** ne sont pas pour les phonèmes mais plutôt pour les syllabes, indique que le nombre de **prototypes** stockés dans le cerveau peut être réduit, avec un minimum de chevauchement pour peu de contextes.

Les propositions de Patricia Kuhl vont dans le sens des catégories phonétiques des voyelles en isolation, qui concernent davantage l'organisation des **prototypes** autour des phonèmes. L'effet des variations contextuelles sur le **prototype** et sur « *The Perceptual Magnet Effect* » semble ne pas être pris en compte dans ses travaux, car :

« [...] *the best exemplars of phonetic categories tend to have more extreme acoustic values than do average productions [...]* ».

(Iverson & Kuhl, 1996 : p. 1136).

Les auteurs suggèrent que les auditeurs préfèrent les instances représentatives d'une catégorie qui maximisent la distinctivité entre les catégories phonémiques, c'est pourquoi les meilleures instances représentatives d'une catégorie phonémique, les **prototypes**, présentent des valeurs acoustiques extrêmes. Les sujets ont tendance à produire des formes réduites de ces cibles extrêmes en parole normale, mais ces mêmes cibles extrêmes peuvent être produites en parole très claire ("*clear speech*"). Ce raisonnement suit la vision de la théorie de *H&H* proposée par Lindblom (1990) : les **prototypes** correspondraient à une forme *Hyper-articulée*, où une maximisation de l'espace acoustique est observée. L'étude établie par Johnson *et al.* (1993) apporte plus d'éclaircissement sur les attentes perceptives des auditeurs et sur l'identité de ces **prototypes**.

2.3.2.7. Le prototype & « The Hyper-Space effect »

Johnson *et al.* (1993) ont montré, sur la base d'expériences perceptives, que la représentation **prototypique** des segments vocaliques d'une langue correspond à un maximum articuloire. Ce maximum articuloire correspond à une forme d'*Hyper-Articulation* de l'espace acoustique des voyelles produites. En effet, les expériences de Johnson explorent les cibles phonétiques en utilisant la méthode d'ajustement de formants, la MOA (pour « *method of adjustment* »), qui révèle les attentes perceptives de l'auditeur des sons de la parole. Dans cette méthode, l'auditeur manipule les paramètres d'un synthétiseur à formants, jusqu'à ce que pour lui la machine prononce correctement le segment vocalique recherché.

Les principaux résultats obtenus par Keith Johnson et ses collègues montrent que l'espace vocalique en perception est de type *Hyper-Articulé* en comparaison avec celui de la production³⁷. La Figure 24 montre les valeurs moyennes des voyelles produites en forme récitée (« *citation form* ») et perçues via la MOA.

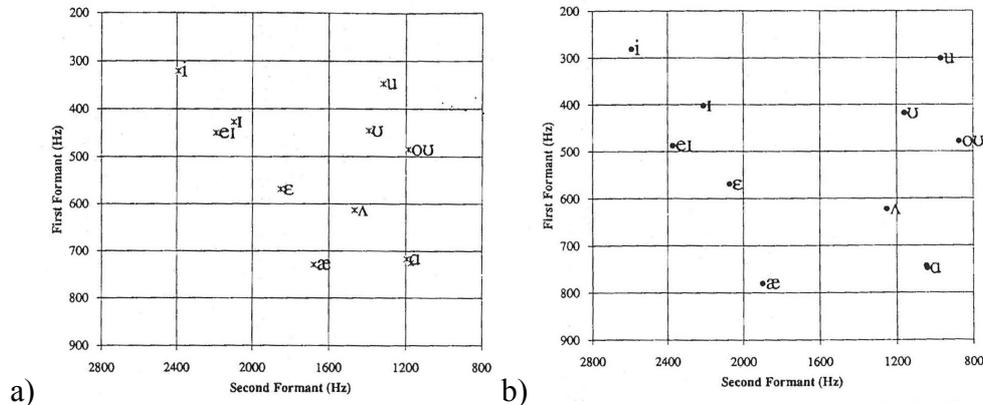


Figure 24 : Moyenne des voyelles de l'anglais américain produites (en a) et perçues via la MOA (en b). On observe « The Hyper-Space effect » en perception (en b). (Emprunté à Johnson *et al.*, 1993).

On observe un « *Hyper-Space effect* » qui affecte l'espace vocalique en perception : les auditeurs choisissent plutôt des voyelles comparables à une production *Hyper-Articulée* qu'à une production *Hypo-Articulée*, suivant ainsi les propos de Lindblom (1990) (dans sa théorie

³⁷ Avec des locuteurs~auditeurs naïfs et non-naïfs.

de *H&H*). Les auditeurs disent de quelle manière le synthétiseur est supposé prononcer un son, en référence à sa représentation en mémoire.

Ainsi, puisque cette tâche est destinée à révéler la nature des cibles phonétiques, les **prototypes**, Johnson *et al.* (1993) interprètent ce résultat en disant que les cibles phonétiques obtenues en perception correspondent à une forme *Hyper-Articulée*, situées dans les limites de l'espace vocalique :

« [...] we found [...] that listeners chose vowels that did not match those produced by any speaker in normal speech. The discrepancy between listener's choices in the MOA and speaker's productions is interesting because it was systematic. The perceptual vowel space was expanded relative to the production space: high vowels were higher, low vowels lower, front vowels more front, and back vowels more back. [...] the MOA vowel space corresponds to the vowel space seen in hyperarticulated speech ».

(Johnson *et al.*, 1993 : p. 509).

En comparant les cibles vocaliques perçues par les auditeurs par la MOA et la production des voyelles en forme *Hyper-Articulée* (« *Hyper-Articulated form* »), les auteurs observent un espace vocalique presque identique (voir Figure 25, partie b). La comparaison entre la production des voyelles en forme *Hyper-Articulée* et normale (« *citation form* ») montre un espace vocalique plus grand dans la première condition par rapport à la seconde, voir Figure 25, partie a.

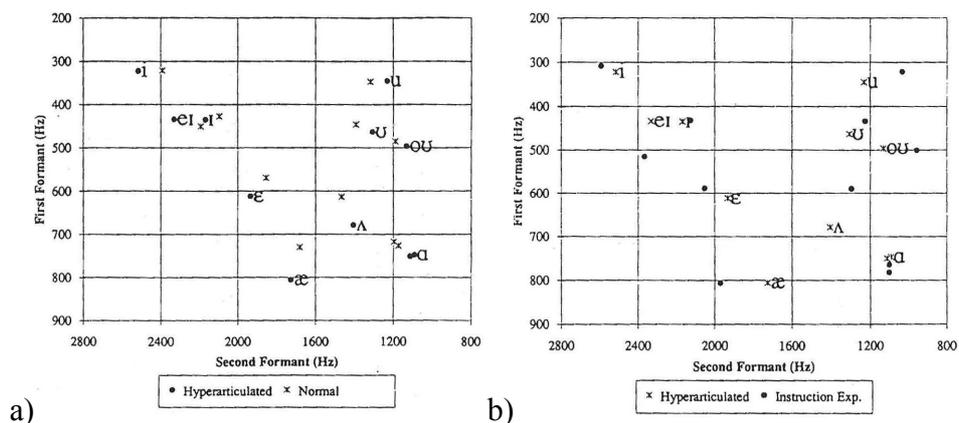


Figure 25 : Moyenne des voyelles de l'anglais américain produites en forme Hyper-Articulée et normale (récitée), (en a)). La figure b montre une comparaison entre les voyelles produites en forme Hyper-Articulée et perçues via la MOA. (Emprunté à Johnson *et al.*, 1993).

Néanmoins quelques réserves peuvent être formulées à propos de ces résultats. « *The Hyper-Space effect* » obtenu est peut-être lié à l'utilisation de stimuli isolés. D'après la théorie *H&H* de Lindblom (1990), plus la quantité d'information est réduite, plus les voyelles sont produites avec une forme *Hyper-Articulée*. On sait qu'en parole normale non contrôlée, les voyelles sont produites en contexte, subissant ainsi des variations **intrinsèques** (des trajectoires dynamiques propres à une voyelle donnée) et/ou **extrinsèques** (coarticulations de type CV, VC ou VV).

Différentes études ont montré que l'utilisation de stimuli synthétiques ou de voix non-familiales affecte la perception des voyelles, qui produit un « *Hyper-Space effect* » important par rapport à celles produites par une voix familière ou une voix humaine resynthétisée (voir Johnson *et al.*, 1999 ; Johnson, 2000). La prise en compte des variations dialectales et régionales, et des différences hommes~femmes modifie également « *The Hyper-Space effect* » obtenu par Johnson *et al.* (1993) (voir les travaux de Whalen *et al.*, 2004_a, 2004_b ; Al-Tamimi & Barkat-Defradas, 2002 ; Barkat-Defradas *et al.*, 2003 ; etc...).

D'une façon générale, les critiques avancées par l'équipe de Douglas Whalen (voir Whalen *et al.*, 2004_a, 2004_b) sur les résultats obtenus par Johnson *et al.* (1993) montrent que « *The Hyper-Space effect* » ne peut être obtenu dans la comparaison entre les perceptions des hommes et des femmes, ou encore entre des auditeurs de différentes origines dialectales. Johnson *et al.* (2004) montrent que même si l'*Hyper-Articulation* des voyelles perçues par la MOA, dans les travaux de Douglas Whalen, n'est pas comparable à celles des voyelles perçues dans le travail de Johnson *et al.* (1993), « *The Hyper-Space effect* » est présent car la taille des espaces vocaliques perçues par la MOA dans les trois études est comparable (i.e., une expansion importante de l'espace vocalique en perception par rapport à celui en production de la parole, même lorsque les voyelles sont produites en forme *Hyper-Articulée*).

Les résultats obtenus dans les différentes études utilisant la MOA peuvent être affectés par un biais méthodologique (voir également les critiques avancées par Whalen *et al.* (2004_b)). Il est possible que le choix des valeurs extrêmes des stimuli synthétiques comme les « meilleures » instances des voyelles par la MOA soit influencé par la qualité des sons

synthétiques qui n'est pas optimale. En effet, ces sons synthétiques ont été générés à partir des valeurs formantiques de l'état stable des voyelles (la cible vocalique).

2.4. CONCLUSIONS

Nous nous sommes intéressé dans ce chapitre à la caractérisation des indices **statiques** et **dynamiques** pertinents en perception de la parole. Nous avons exposé les différentes approches en perception des voyelles et en particulier les approches s'appuyant sur l'utilisation des indices **statiques** (les approches « *Simple Target* »), ou **dynamiques** (les approches « *Elaborated Target* » et « *Dynamic Specification* »). Comme nous l'avons vu, ces différentes approches cherchent à identifier la façon la plus simple pour caractériser les invariants acoustiques, à partir desquels des tests de perception permettent d'appréhender leur rôle dans la perception de la parole. Pour les partisans des approches « *Simple Target* » ce sont des informations de la cible vocalique, ou des indices **statiques**, qui sont suffisantes à l'identification des voyelles. Pour les deux autres approches en perception des voyelles : l'« *Elaborated Target* » et le « *Dynamic Specification* » ce sont différents indices supplémentaires et/ou **dynamiques** à l'intérieur des réalisations vocaliques qui facilitent la discrimination des voyelles.

Nous avons ensuite présenté les différentes théories développées en perception de la parole, i.e., comme processus général, ou comme processus spécifique. Nous avons distingué les approches passives et actives. Pour les premières, la relation locuteur~auditeur n'a pas de rôle important dans les mécanismes de perception de la parole et plus particulièrement des voyelles. Pour les secondes, la perception de la parole est un processus spécifique qui fait intervenir le locuteur~auditeur dans l'identification des voyelles et des consonnes, car l'auditeur fait usage des connaissances qu'il a de son système de production : la production fait partie intégrante de la perception de la parole.

Nous présentons dans la partie qui suit la problématique du travail.

Chapitre 3.

PROBLÉMATIQUE

Cette thèse a pour double objectif, d'une part, de comprendre l'impact des indices **statiques** et **dynamiques** en perception des voyelles de l'arabe et d'autre part de comparer le rôle de la densité des systèmes vocaliques sur la dispersion des voyelles en arabe et en français.

Concrètement, nous cherchons à trouver la meilleure description de la voyelle en arabe ; description expliquant sa perception par les auditeurs arabophones. Toutes les études présentées dans la partie théorique nous conduisent à nous questionner sur la représentation des voyelles de l'arabe : vue l'impossibilité de produire la voyelle en arabe en isolation, est-ce que sa représentation suivant le modèle « *Simple Target* » est adaptée ? Une description des voyelles en termes **dynamiques** (i.e. en suivant l'un des deux modèles, le modèle « *Elaborated Target* » ou « *Dynamic Specification* ») contribue-t-elle à une meilleure appréhension du statut de la voyelle en arabe ?

En bref, quel est le rôle des indices acoustiques **dynamiques** observés sur des données de production des voyelles en arabe pour décrire les systèmes vocaliques ? Quelle est l'incidence de ces mêmes indices dans la description du système vocalique du français ? Nous chercherons tout d'abord à identifier les différents traits **statiques** et **dynamiques** à partir des données de production des voyelles en arabe jordanien et marocain et en français. Une fois ces traits relevés, des tests en perception permettront de mesurer leur importance dans les langues étudiées. À partir des descriptions des voyelles en arabe et en français, nous nous attendons à observer plus de facilité dans la perception des voyelles en arabe en comparaison avec le français lorsque des indices **dynamiques** sont disponibles. Puisque les contraintes phonotactiques de l'arabe font que les voyelles ne peuvent être produites qu'en association avec les consonnes adjacentes, une représentation **dynamique** des voyelles permettrait leur

identification par les arabophones. L'utilisation de ces mêmes indices en français ne devrait pas notablement faciliter l'identification des voyelles françaises, car la voyelle en français peut être produite en isolation, à cause des contraintes phonotactiques différentes. À partir des prédictions proposées par René Carré (voir Carré, 2004, 2007 (à paraître) ; Carré *et al.*, 2004), nous pouvons supposer que les indices **statiques** et **dynamiques** soient toujours mis à profit par les auditeurs de différentes langues : ces deux informations apportent autant de précision et de facilité à l'identification des voyelles.

Nous avons choisi de comparer l'arabe et le français non seulement à cause des différences morphologiques et phonotactiques, mais surtout à cause de la différence du nombre de voyelles dans chaque langue. Rappelons que ces langues comportent cinq, huit et onze voyelles orales (pour l'arabe marocain, l'arabe jordanien et le français, respectivement), situation qui doit nous permettre d'appréhender le statut de celles-ci et de tester ensuite l'impact de la densité des systèmes vocaliques sur la dispersion des voyelles.

EFFETS DE LA DENSITÉ DES SYSTÈMES

Par densité des systèmes, nous entendons le nombre de segments vocaliques, ou consonantiques, qu'une langue possède. Au niveau vocalique, les langues diffèrent par le nombre et l'organisation de leurs segments : de trois à plus de vingt voyelles distinctes (Maddieson, 1984 ; Maddieson & Precoda, 1991 ; Vallée, 1994). Différentes études typologiques ont montré que cette organisation est régie par certaines contraintes articulatoires et perceptives et qu'elle est fortement corrélée à la densité (i.e. nombre de voyelles) du système vocalique (Maddieson, 1984 ; Vallée, 1994). En se basant sur ces travaux translinguistiques, Stevens (1972, 1989, 2003) et Lindblom (1986) ont étudié comment le nombre de segments vocaliques dans une langue affecte leur dispersion à l'intérieur de l'espace vocalique. Différentes théories visant à prédire l'organisation des systèmes vocaliques ont ainsi vu le jour. Ces différentes théories, que nous rappelons brièvement, cherchent à évaluer différents principes universels au niveau de l'organisation des espaces vocaliques à travers les langues.

Basée sur la relation non-linéaire entre les mouvements articulatoires et leurs corrélats acoustiques, la **théorie quantique (TQ)** proposée par Stevens (1972, 1989, 2003) explique l'existence de certaines régions de stabilité dans l'espace phonétique correspondant aux voyelles extrêmes /i/, /a/ et /u/ qu'il nomme « *Hot-Spots* ». Ces « *Hot-Spots* » doivent se positionner approximativement dans la même zone à travers les langues et indépendamment de la taille de leur inventaire vocalique. De plus, puisque ces « *Hot-Spots* » se situent approximativement dans des zones de stabilité phonétique, ils présenteraient moins de variabilité intra-catégorie que les voyelles non extrêmes (voir également, § 2.3.2.3).

À l'inverse, les différentes théories proposées par Björn Lindblom (Liljencrants & Lindblom, 1972 ; Lindblom, 1986) propose que la dispersion des voyelles dans les espaces vocaliques doit être **maximale** et **adaptée**. Dans la première version de cette théorie : **la théorie de la dispersion maximale (TDM)**, Liljencrants & Lindblom (1972) expliquent que pour qu'un système vocalique soit optimal, la distance entre les unités qui le constituent, doit être **maximale**. Par la suite, Lindblom (1986) a amélioré son modèle de prédictions des voyelles en proposant qu'un système vocalique requiert une dispersion non pas **maximale** entre les unités qui le constituent, mais une dispersion **adaptée** d'où le nom de ce modèle : **la théorie de la dispersion adaptée (TDA)**. L'auteur explique que l'organisation des sons de la parole est gouvernée par une **dispersion adaptée** de leurs éléments selon le **principe de contraste perceptuel suffisant**. D'après cette théorie, les voyelles d'une langue sont organisées de telle manière à être perceptuellement distinctes les unes des autres. Ainsi, les langues ayant des espaces vocaliques à faible densité peuvent présenter une grande variabilité intra-catégorie par rapport à celles à forte densité.

À partir des travaux de Björn Lindblom sur ses théories **TDM** et **TDA**, une nouvelle théorie de prédiction des voyelles dans les langues de monde a vu le jour : **la théorie de la dispersion-focalisation (TDF)** (Schwartz *et al.*, 1989 ; Vallée, 1994 ; Boë *et al.*, 1996 ; Schwartz *et al.*, 1997_a, 1997_b ; etc.). Schwartz *et al.* (1997_b) expliquent que les systèmes vocaliques sont régis par deux contraintes qui sont en interdépendance : la **dispersion** (ou la distance entre les voyelles), qui augmente la distance maximale entre les voyelles et la **focalisation** (ou la saillance spectrale à l'intérieure de chaque voyelle), qui met en avant le

rôle du formant « *effectif* » pour une voyelle donnée (voir le modèle « *Simple Target* », Chapitre 2). Dans cette théorie, les voyelles sont représentées sur la base de F_1 et de F' : les voyelles postérieures sont perçues en fonction de leur **centre de gravité** résultant du rapprochement entre de F_1 & F_2 , ou F_1' ; les antérieures, en fonction de F_1 et F_2' (F_2' = le **centre de gravité** de F_2 et F_3 , ou F_3 et F_4 pour /i/). Ce sont alors deux facteurs intra- et inter-voyelles qui conditionnent l'organisation des systèmes vocaliques : les voyelles dans un système doivent être **distinctes** et **focales** en même temps. Ces facteurs peuvent expliquer en partie les effets de la densité des systèmes vocaliques sur la stabilité ou la réduction de l'espace vocalique, car la **dispersion** explique la distance **maximale** et **suffisante** entre les voyelles d'un système ; et la **focalisation** peut expliquer la dispersion intra-catégorie qui doit être toujours faible : l'augmentation ou la diminution de l'énergie responsable de la **focalisation** produit une instabilité dans le système car les points représentant les voyelles « optimales » seront très rapprochés et vont ainsi créer une confusion dans le système. C'est pourquoi nous pouvons supposer que d'après la **TDF**, la dispersion intra-catégorie doit être toujours constante entre les différentes voyelles pour éviter des confusions entre elles.

En prenant en considération le fait que l'organisation des espaces vocaliques est fortement corrélée à la densité des systèmes, on peut proposer différents schémas d'organisation des espaces vocaliques en fonction du nombre de segments. Si l'on prend deux langues différant dans leur nombre de segments vocaliques : une langue A avec dix voyelles et une langue B avec trois voyelles, on obtient quatre cas de figures³⁸ (voir Figure 26) :

- Cas 1 L'espace vocalique reste constant pour les deux langues, indépendamment de la densité des systèmes, où l'espace articuloire et acoustique entre deux voyelles reste inexploité, suivant ainsi la stabilité proposée par la **TQ** et la **focalisation** proposée par la **TDF**,
- Cas 2 L'espace vocalique subit l'effet de la densité des systèmes, autrement dit, l'organisation des voyelles est fonction de la densité des systèmes ; plus une langue a de voyelles, plus l'espace vocalique est grand. La dispersion intra-catégorie vocalique sera identique entre les deux langues. L'espace articuloire et acoustique

³⁸ Les trois premiers cas sont empruntés à Meunier *et al.* (2003).

entre deux voyelles reste inexploité, suivant ainsi partiellement les propositions de la **TDA**,

- Cas 3 L'espace vocalique reste constant pour les deux langues, indépendamment de la densité des systèmes, i.e. les trois voyelles extrêmes formeront des « *Hot-Spots* », avec une grande dispersion intra-catégorie vocalique, couvrant ainsi la totalité de l'espace articulatoire et acoustique, donc un mélange de propositions de la **TQ**, de la **TDA** et de la **TDF** et enfin
- Cas 4 L'espace vocalique subit l'effet de la densité des systèmes et l'espace vocalique de la langue B sera plus réduit que celui de la langue A. Une grande variabilité intra-catégorie vocalique sera observée dans la langue B par rapport à la langue A, suivant ainsi totalement les prévisions de la **TDA**.

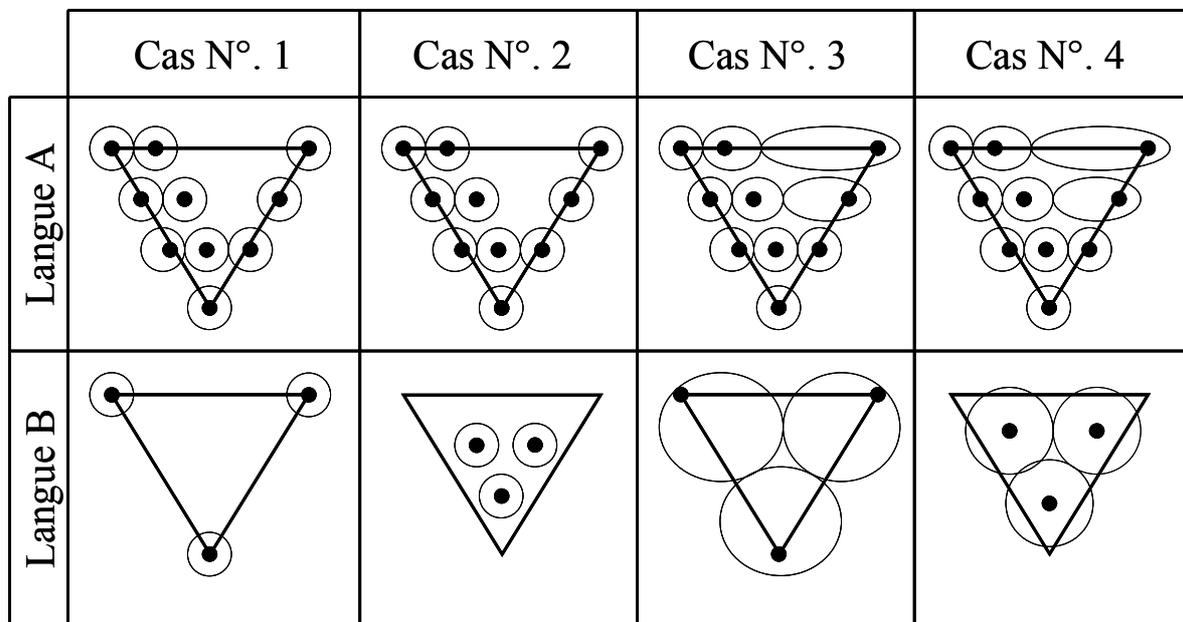


Figure 26 : Schémas hypothétiques des quatre cas de figures de l'organisation de l'espace vocalique de la langue A avec dix voyelles et la langue B avec trois voyelles. (Inspiré en partie de Meunier *et al.*, 2003).

Ces quatre cas de figures sont des prédictions théoriques concernant l'organisation des espaces vocaliques de deux types de langues extrêmes ; une langue à trois voyelles et une autre à dix. Quelques études ont déjà confirmé partiellement leur validité (voir entres autres : Bradlow, 1995, sur l'anglais (onze monophthongues) et l'espagnol (cinq voyelles) ; Engstrand & Krull, 1991, sur sept langues (systèmes vocaliques qui varient de cinq à dix-sept segments) ; Livijn, 2000, sur vingt-huit langues (systèmes vocaliques qui varient de quatre à quatorze segments) ; Meunier *et al.*, 2003, sur l'anglais (dix monophthongues), le français (onze voyelles orales) et l'espagnol (cinq voyelles)).

Nous examinerons dans notre travail les effets de la densité des systèmes vocaliques sur la dispersion des espaces vocaliques et des voyelles en arabe marocain, en arabe jordanien et en français. Nous évaluerons nos résultats selon les prédictions de la **TQ**, de la **TDA** et de la **TDF**.

Section II. CADRE EXPÉRIMENTAL

Chapitre 4.

RÉSULTATS EN PRODUCTION DES VOYELLES

4.1. INTRODUCTION

Cette section, composée de deux chapitres (chapitres N°. 4 & 5), s'intéresse aux aspects expérimentaux de notre travail. Les résultats obtenus en production des voyelles seront développés dans le Chapitre 4 ; ceux obtenus en perception, dans le Chapitre 5. En guise d'introduction de ce chapitre, nous présentons une description du corpus, des différents locuteurs participant à cette étude, ainsi que des protocoles expérimentaux élaborés pour tester nos hypothèses. Les résultats en production des voyelles seront développés en traitant des aspects i) **statiques** (i.e., les « cibles vocaliques » correspondant aux valeurs formantiques à l'état stable) et ii) **dynamiques** (i.e., les indices **extrinsèques** (voir § 2.2.3) correspondant aux influences consonantiques sur les voyelles obtenues à partir d'analyses de régressions linéaires et polynomiales). Dans chaque sous-partie, nous nous proposons d'évaluer le rôle de la densité des systèmes vocaliques sur la dispersion des voyelles de chaque langue. Ainsi, du point de vue des traits **statiques** des voyelles, nous évaluerons la taille des aires de dispersion des voyelles de chaque système pour les trois voyelles /i a u/, car ce sont les seules voyelles *phonologiquement* communes entre les trois langues et par ailleurs, ce sont les voyelles qui délimitent l'espace vocalique. En ce qui concerne les traits **dynamiques**, nous évaluerons le rôle de la densité des systèmes sur les **pentés formantiques** pour ces mêmes trois voyelles /i a u/. Nous terminons ce chapitre par une présentation des spécificités de chaque langue notamment le statut des voyelles brèves en arabe et celui des voyelles dites « médianes » ainsi que l'opposition /a ɑ/ en français. Nous présenterons également dans chaque sous-partie les résultats d'une méthode de classification des voyelles qui nous permettrons de vérifier d'une

part, la validité des effets de la densité des systèmes sur les analyses utilisés, et d'autre part, les spécificités de chaque langue.

4.2. MÉTHODOLOGIE DU TRAVAIL

4.2.1. Locuteurs

Trente locuteurs masculins (10 locuteurs par langue) ont participé aux tests de production des voyelles. Les locuteurs jordaniens sont originaires de la région d'Irbid, les marocains de Mohammedia (faisant partie du Grand Casablanca) et les français de Lyon et de Toulouse (cinq locuteurs pour chaque ville)³⁹. Ils ne présentent aucun trouble du langage, ni articulaire ni auditif (testés en audiométrie). L'intervalle d'âge est situé entre 18 et 30 ans. La situation socioprofessionnelle de tous les locuteurs est homogène, à savoir des étudiants à l'université, issus d'un même milieu socioprofessionnel et vivant tous en ville. Nous avons essayé de choisir des locuteurs monolingues, mais la tâche s'est avérée difficile : les jordaniens, dont certains sont d'origine palestinienne, sont plus ou moins bilingues (avec soit le français, soit l'anglais), les marocains, non berbérophones, ont des connaissances plus ou moins approfondies du français, tandis que les français ont quelques connaissances d'autres langues indo-européennes. Tous les locuteurs arabophones ont approximativement le même niveau de connaissance de l'arabe standard, (voir § 1.3.2) : plus spécifiquement, les locuteurs jordaniens ont suivi, tout au long de leur formation universitaire, des cours en arabe standard et sont considérés comme « bilingues » arabe dialectal/arabe standard ; les marocains quant à eux, sont plus au moins « bilingues » car ils n'ont pas suivie de formations approfondies en arabe standard.

³⁹ Le corpus dont nous disposons en français ne présente que cinq locuteurs natifs de chaque région. Les différences existantes au niveau vocalique entre les locuteurs de la région du Centre et du Sud de la France concernent le traitement des voyelles intermédiaires, qui sont phonologiquement distinctes pour les locuteurs de la région Centre et sont complémentaires pour les locuteurs de la région Sud. Dans ce travail, nous n'étudions pas cette différence ni en production ni en perception de la parole. Nous cherchons à comparer le système vocalique du français en tant qu'unité globale, avec les systèmes vocaliques de l'arabe jordanien et de l'arabe marocain.

4.2.2. Corpus

Trois corpus ont été élaborés pour chaque langue, et présentent les mêmes structures : C_1V_1 , $C_1V_1C_2$, C_1V_1C , C_1V_1CC , $C_1V_1C_2C$, C_1V_1CV , $C_1V_1C_2VC$, C_1V_1CVC , C_1V_1CCV , CVC_1V_1C , C_1V_1CVCV , $CVCC_1V_1C_2$, C_1V_1CCVC , $C_1V_1C_2CVC$, ou VC_1V_1C , où C_1 ou C_2 est l'une des trois consonnes phonologiquement communes entre l'arabe marocain, l'arabe jordanien et le français (AM, AJ et FR désormais), soit /b d k/ (et /d^ʕ/ en AM et AJ), et V_1 chaque voyelle. Nous n'avons pas pu constituer des corpus de paires minimales pour chacune des langues, par conséquent nous avons essayé, autant que possible, de les présenter dans des contextes adéquats (voir Tableaux 3 à 5).

Tableau 3 : Liste d'items utilisés pour la tâche de production en AJ.

	b	sens	d	sens	d ^ʕ	sens	k	sens
i:	bir	puits	di:nak	ta religion	mad ^ʕ i:q	étroit	ki:s	sac
i	bint	fille	dijja	indemnité	d ^ʕ id ^ʕ	contre	kilma	mot
e:	bet	maison	de:r	monastère	d ^ʕ e:f	invité	ke:f	comment ?
a:	ba:b	porte	da:r	maison	d ^ʕ a:q	rétrécit	ka:f	lettre « k »
a	bard	froid	dam	sang	d ^ʕ abt ^ʕ	exactement	kaf	paume
o:	bo:se	bisou	do:ri:	mon tour	d ^ʕ aw/d ^ʕ o:ʔ	lumière	ko:n	univers
u:	bu:ma	chouette	du:d	vers (de terre)	maʕd ^ʕ u:d ^ʕ	mordu	ku:sa:	courgette
u	burj	pont	dub	ours	d ^ʕ uha	avant-midi	kul	chacun

Tableau 4 : Liste d'items utilisés pour la tâche de production en AM.

	b	sens	d	sens	d ^ʕ	sens	k	sens
i:	bir	puits	di:b	loup	d ^ʕ i:f	invité	ki:f	comment ?
ə	berd	froid	demli:ʒ	bracelet	d ^ʕ ebt ^ʕ	exactement	kefta	boulette de viande
a:	ba:t	resté	da:b	fondu	d ^ʕ a:q	rétrécit	ka:fi	suffisant
ʊ	buxxar	encens	dul	humiliation	d ^ʕ ʊlma	obscurité	kul	chacun
u:	bu:l	pisse	du:da	vers (de terre)	d ^ʕ u:ss	le deux	ku:mi:ra	voiture

Tableau 5 : Liste d'items utilisés pour la tâche de production en FR.

	b	Transcription	d	Transcription	k	Transcription
i	bilan	bilā	dis	di	qui	ki
e	bécasse	bekas	dé	de	képi	kepi
ɛ	baie	bɛ	dès	dɛ	caisson	kesō
a	bal	bal	date	dat	cale	kal
ɑ	basse	bas	d'âme	dɑm	case	kɑz
ɔ	bol	bɔl	dors	dɔʁ	cote	kɔt
o	Baule	bol	dôme	dɔm	côte	kot
u	bout	bu	doux	du	coup	ku
œ	bœuf	bœf	odeur	ɔdœʁ	cœur	kœʁ
ø	bœufs	bø	vendeuse	vādøz	aqueuse	akøz
y	bu	by	du	dy	cul	ky

4.2.3. *Protocole expérimental en production*

Les expériences pour les trois langues ont été réalisées via un logiciel conçu spécialement pour cette étude par François Pellegrino en Tcl/Tk (version 8.4.3.0, 2003) et Snack Sound ToolKit (version 2.2 ; Sjölander, 2002), avec un affichage en Unicode. L'interface du programme permet au locuteur de lire les consignes⁴⁰ et de s'identifier (un code à 5 caractères composés du code de la langue (AM, AJ ou FR), du symbole M pour masculin et du numéro de l'auditeur en deux chiffres, e.g., AJM03). Une fois identifié, le locuteur doit enregistrer une phrase permettant de régler l'intensité de sa production (la phrase en FR était « *vous arrivez plus tôt demain* » ; deux autres phrases ont été proposées en arabe). Le microphone utilisé est le Sony MS 907 et est placé à 30 cm de distance de la bouche du locuteur. Il est demandé au locuteur de ne pas modifier, dans la mesure du possible, la distance entre lui et le micro. Nous procédons à l'enregistrement d'un ton de référence qui permet de calibrer l'intensité de la production des locuteurs. Pour ce faire, le sujet place près de sa bouche et en direction du microphone un émetteur qui produit un « la » (i.e., un son pur à 440 Hz). L'enregistrement de ce ton permet d'étalonner l'intensité du signal durant l'expérience dans le cas où le locuteur modifie sa position par rapport au microphone. Nous

⁴⁰ Dans la mesure du possible, les consignes écrites en français, ont été expliquées à plusieurs reprises aux locuteurs arabophones au départ en français, et si besoin en arabe dialectal par un locuteur natif jordanien (nous-même) et marocain (Thami Benkirane).

avons utilisé également un amplificateur du signal, car l'intensité de certains locuteurs ne permettait pas d'obtenir un signal acoustique de bonne qualité.

Après une phase de familiarisation et d'entraînement des locuteurs avec l'interface du logiciel, une présentation des items est alors proposée à l'intérieur d'une phrase porteuse adaptée à chaque mot, aléatoirement avec cinq répétitions. La tâche des locuteurs consiste à produire les voyelles dans le Mot présenté, dans la Syllabe qui les contient et enfin en Isolation⁴¹, sans produire la phrase porteuse, en essayant de conserver la qualité de production de la voyelle en mot (e.g., [bo:se~bo:~o:] = "bisou" en AJ), à débit moyen et avec un style de production non-marqué (voir Figure 27 pour visualiser l'interface du logiciel). Pour chaque langue, nous voulons évaluer les différences entre les réalisations des voyelles produites dans les trois situations de production en Mot, Syllabe et en Isolation afin de mettre en évidence les effets des différences contextuelles sur le degré plus ou moins invariant des voyelles (*H&H*, Lindblom, 1990). Les enregistrements ont été effectués dans une chambre insonorisée⁴² ; et numérisés directement sur un PC avec un taux d'échantillonnage de 22 KHz, 16 bits, mono. La durée de passation de cette tâche variait entre 30 min et 1 heure par locuteur. Ainsi nous obtenons un corpus très riche contenant l'équivalent de 15 à 30h de parole.

La tâche de production a été adaptée aux deux dialectes arabes, qui sont des dialectes de tradition orale. Le système d'écriture de l'arabe standard moderne a été utilisé pour présenter visuellement les items. Ce type de présentation visuelle nous a permis de réduire au maximum les influences de prononciation que peuvent avoir les expérimentateurs sur les locuteurs. En arabe, seules les voyelles longues sont transcrites avec des lettres de l'alphabet ; les brèves quant à elles, sont marquées par des symboles suscrits ou souscrits. Ainsi nous avons décidé de présenter tous les items arabes sans « vocalisations » afin de réduire au maximum l'influence de l'écrit et de l'arabe standard sur la prononciation des locuteurs. La phrase porteuse a donc un rôle essentiel pour déterminer le sens des mots.

⁴¹ Nous utiliserons désormais le terme « **situation de production** » pour parler des trois contextes de réalisation des voyelles, i.e., en Mot, en Syllabe et en Isolation.

⁴² Ou à défaut dans un endroit calme.

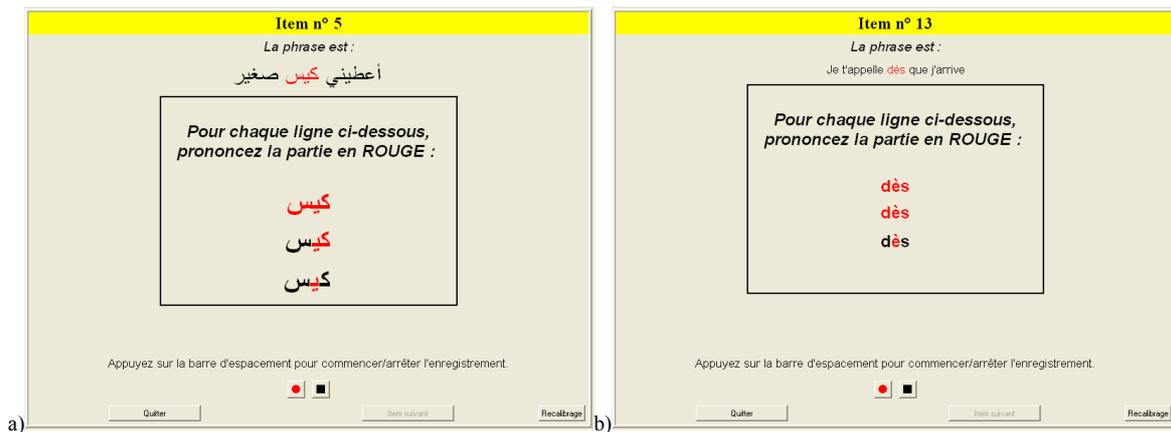


Figure 27 : Interfaces du logiciel dans la tâche d'enregistrement des voyelles en AJ (partie a) et en FR (partie b), présentant la phrase porteuse avec l'item à produire. Les parties que les locuteurs doivent produire en fonction des situations de production apparaissent en rouge.

4.2.4. Segmentation du corpus

Tous les items produits par les locuteurs des trois langues, soit 4100 items (un item = un fichier enregistré, donc on obtient trois voyelles différentes par fichier), ont été segmentés manuellement à partir de Praat© (version 4.4.34. ; Boersma & Weenink, 2006). Chaque fichier son contient la production des trois réalisations de la voyelle en fonction des trois situations de production. La Figure 28 montre un exemple de segmentation du mot [di:nak] « ta religion » en AJ. Trois niveaux de segmentation ont été utilisés : le premier concerne chaque situation de production : Mot, Syllabe et Isolation, appelé « stimulus » ; le deuxième concerne la barre d'explosion et la voyelle, et appelé « syllabe » ; et enfin le troisième concerne la voyelle elle-même et ainsi appelé « voyelle ». Les conventions de segmentation de chaque niveau étaient les suivantes :

- Pour le premier niveau : « stimulus », le début et la fin de chaque situation de production sont déterminés par l'apparition/disparition d'onde dans la première/dernière période sur le signal acoustique en fonction de la modification brusque de l'intensité et de f_0 ;
- Pour le deuxième niveau : « syllabe », la frontière de début de segment est insérée au début de la barre d'explosion pour les situations Mot et Syllabe, et au début de la

voyelle pour la situation Isolation ; la frontière de fin est insérée à la fin de la dernière période de la voyelle au « zero crossing » de celle-ci⁴³ ; et enfin

- Pour le troisième niveau : « voyelle », la frontière de début/fin de la voyelle est insérée au début de la première/dernière période de chaque voyelle au « zero crossing ». Ces frontières correspondent ainsi à l'**onset/offset** pour chaque voyelle et servent pour le calcul de la durée vocalique.

Pour les trois niveaux de segmentation, nous avons utilisé les paramètres de l'intensité et de f_0 pour déterminer le début et la fin de chaque partie du signal. Ainsi lorsque la valeur de l'intensité et de f_0 est supérieure à 0 et/ou le changement de valeurs de ces paramètres est brusque, la frontière est insérée. La segmentation du deuxième niveau a été proposée dans le but d'effectuer les analyses des **équations de Locus** et de les comparer avec les résultats des **pentés formantiques** (voir § 4.3.2.1).

Un « script » spécifique a été écrit sous Praat pour accélérer la phase de segmentation du corpus, qui a duré près de 4 mois pour obtenir une segmentation très fine de tous les items. Tous les fichiers de segmentation ont été enregistrés sous TextGrid (format de Praat).

⁴³ Le « zero crossing » correspond au point où le milieu de la période entre en croisement avec le point zéro du signal acoustique.

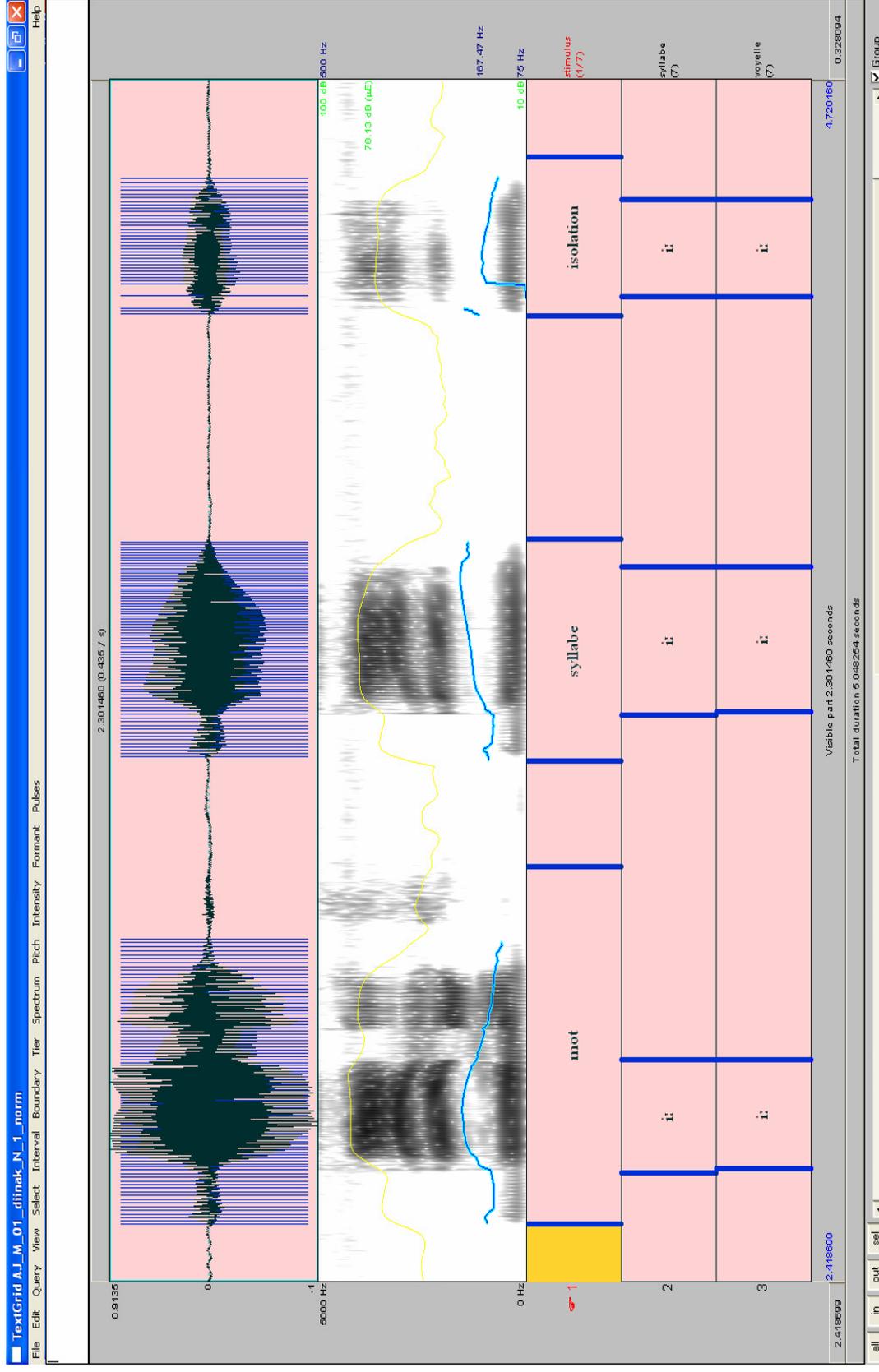


Figure 28 : Segmentation du mot [di:nak] « ta religion » en AJ, produit par le locuteur AJM01 dans les trois situations de production.

Le corpus obtenu est conséquent, ce qui nous permettra de proposer une description approfondie des systèmes vocaliques des trois systèmes étudiés. Nous présentons dans les parties qui suivent les résultats obtenus en fonction des analyses **statiques** et **dynamiques**, en allant du plus général au plus spécifique, et procédons ensuite à une comparaison de la pertinence des indices utilisés dans la description des trois langues.

4.3. RÉSULTATS INTER-LANGUES

Nous avons expliqué que la densité des systèmes affecte la dispersion des voyelles. Les différentes prédictions théoriques formulées dans les parties précédentes devraient expliquer certaines tendances observées dans les langues de monde. Selon la **théorie quantique**, les trois voyelles /i a u/ sont stables dans toutes les langues et subissent le moins les effets de la densité des systèmes, i.e., nous ne devons observer aucun changement de leur qualité selon la densité des systèmes. Selon la **théorie de la dispersion adaptée**, ces mêmes voyelles doivent présenter une dispersion perceptuellement suffisante pour les distinguer : i.e., leur qualité peut subir des différences en fonction de la densité des systèmes. Enfin, pour la **théorie de la dispersion-focalisation**, ces trois voyelles vont présenter une dispersion perceptuellement suffisante et en même temps, elles sont focales, i.e., cet aspect de **focalisation** va rendre leur dispersion moins marquée : on se rapprochera davantage des prédictions de la **théorie quantique**.

Nous nous concentrerons donc sur l'étude des effets de la densité des systèmes sur la position et la taille des aires de dispersion des trois voyelles /i a u/ (voyelles longues en arabe) produites dans les trois langues. Nous avons choisi de ne comparer que ces trois voyelles car ce sont les seules voyelles *phonologiquement* communes entre les trois langues et ce sont celles qui délimitent l'espace vocalique. Nous voulons ainsi évaluer la stabilité de ces voyelles dans les trois langues étudiées en fonction de la densité des systèmes. Nous savons néanmoins, que l'espace vocalique dans chaque langue est constitué d'autres voyelles, ainsi, et lorsqu'il est nécessaire, nous évaluerons l'apport des autres voyelles sur les effets de la densité des systèmes.

Nous évaluerons ainsi dans les parties qui suivent les effets de la densité des systèmes vocaliques sur les voyelles de chaque langue en fonction des deux analyses proposées. Nous commençons par examiner ces effets en fonction des traits **statiques** et **dynamiques** (désormais nous emploierons les termes en **statique** pour parler des traits **statiques** (obtenus par les cibles vocaliques et en **dynamique** pour parler des traits **dynamiques** obtenus par les **pentés formantiques**). Nous terminons cette partie par une présentation des résultats d'une méthode de classification des voyelles, i.e., l'analyse discriminante. Dans chaque sous-partie, nous présenterons les méthodes d'analyse et de traitement que nous avons appliquées à nos données et présenterons les résultats obtenus en fonction de chaque trait. À la fin, nous procéderons à une comparaison de la pertinence des traits utilisés.

4.3.1. Analyses statiques des formants des voyelles

4.3.1.1. Analyse et traitement des données

Notre corpus contient 12036 voyelles qui ont été produites par les locuteurs des trois langues (2952 voyelles en AM, 4286 en AJ et 4798 en FR)⁴⁴. Les mesures acoustiques de F_1 , F_2 & F_3 ont été effectuées au milieu temporel de chaque voyelle (i.e., 50% de la durée vocalique⁴⁵), comme proposées dans la littérature, en utilisant l'algorithme d'extraction de

⁴⁴ En AJ, le corpus de départ devait contenir 4800 voyelles. Nous obtenons cette différence à cause de la réalisation variable de la voyelle /o:/ en contexte /d^h/ tantôt en monophongue tantôt en diphtongue par les locuteurs jordaniens ; de problèmes techniques liés à la qualité très détérioré du signal acoustique obtenu pour les voyelles /i u/ en contexte vélaire ; et la non réalisation de certains items par certains locuteurs, ainsi le nombre de voyelles réellement analysées est de 4286. En FR, la voyelle /ø/ en contexte bilabial n'a pas été réalisée par les locuteurs français à des mêmes problèmes de détérioration du signal acoustique obtenu.

⁴⁵ L'**onset** vocalique correspond au début de la voyelle et est obtenu à partir de la première période de la voyelle après le relâchement de la barre d'explosion et correspond ainsi au temps 0 ms ; l'**offset** vocalique correspond à la fin de la voyelle et est obtenu à partir de la dernière période de la voyelle ; la **durée vocalique** est déterminée sur la base de la différence du « temps » obtenue entre l'**onset** et l'**offset** pour chaque voyelle.

formants « Burg »⁴⁶ proposé dans Praat (analyse LPC par auto-corrélation, avec 24 coefficients) avec une fenêtre d'analyse gaussienne de 12,5 ms et un pas de déplacement de 5 ms. Les valeurs formantiques ont été vérifiées manuellement afin de corriger les possibles erreurs d'extraction automatique. Une fois les valeurs vérifiées, elles ont été converties ensuite en Bark (suivant la formule : $F_{Bark} = 7 * ASINH(F_{Hertz}/650)$ ⁴⁷ proposée par Schroeder *et al.*, 1979) pour procéder à une normalisation de données entre les locuteurs.

Afin de mesurer l'effet de la situation de production et du lieu d'articulation sur la dispersion des voyelles et des catégories vocaliques, nous avons calculé la taille des aires de dispersion en utilisant la méthode de calcul de l'Enveloppe Convexe, proposée dans MATLAB® (version 6.5 ; The MathWorks, 2002).

Le principe de base de la méthode de calcul de l'enveloppe convexe est d'obtenir le plus petit polygone convexe contenant tous les points présents dans cet espace. Ainsi, en partant d'une représentation bidimensionnelle des voyelles (i.e., un nuage de point obtenu à partir des valeurs formantiques de F_1 & F_2), l'algorithme va calculer les différents triangles définis par trois points appartenant au nuage de points qu'on peut obtenir et ainsi l'enveloppe convexe va coïncider avec la fusion de tous les triangles obtenus (voir Figure 29). Une fois les différents triangles fusionnés, on obtient un polygone fermé qui définit l'enveloppe qui suit le contour « intuitif » représenté par tous les points situés dans les limites de l'espace.

⁴⁶ Nous avons utilisé plusieurs méthodes d'extraction de formants basées sur la méthode *Linear Predicting Coding* par auto-corrélation. À chaque fois, nous changeons le nombre de coefficient LPC en partant de 22 coefficients jusqu'à 26. Nous nous sommes basés sur les données de la littérature et surtout sur les propositions de Ladefoged (1996) pour choisir le coefficient le plus pertinent. Ainsi, pour Ladefoged (1996), le nombre de coefficient doit être calculé à partir du taux d'échantillonnage auquel on ajoute ± 2 : dans notre cas, le taux obtenu était de 22KHz, donc on obtient un coefficient de 22, à cela on doit ajouter ± 2 . Pour certains locuteurs, le coefficient idéal était de 23 ; pour d'autres 25. Nous avons comparé les résultats obtenus avec les différents coefficients avec ceux obtenus à partir de la méthode « Burg » qui consiste à limiter le nombre de formants à obtenir à 5 dans un intervalle de 0 à 5000 Hz pour un homme (et 5500 Hz pour une femme). Les résultats ont montré que c'est avec la méthode « Burg » que les mesures obtenues contiennent peu d'erreur d'extraction.

⁴⁷ Où *ASINH* renvoie au Sinus hyperbolique inverse d'un nombre.

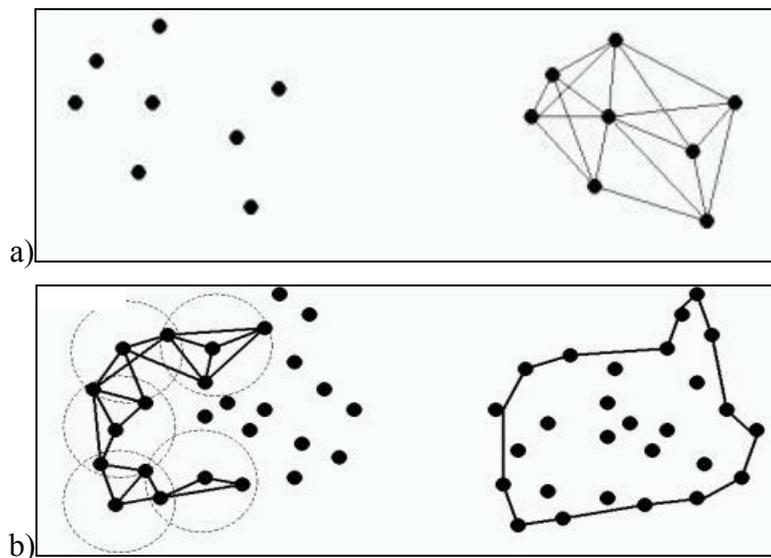


Figure 29 : Schémas montrant le principe de base du calcul de l'enveloppe convexe. En partant de plusieurs points (partie a), on calcul les différents triangles possibles, on les fusionne et on obtient un polygone (partie b).

Une fois l'enveloppe convexe, obtenue, l'aire de ce polygone fermé est obtenue à partir de l'algorithme de calcul « polyarea » proposé dans MATLAB. Ainsi on mesure la grandeur de la surface représentant le polygone. L'aire obtenue est exprimée au carré (en fonction de l'unité de mesure). Dans notre cas, nous obtenons la plus petite surface qui enveloppe les voyelles dans un espace acoustique en Bark, donc l'aire obtenue sera exprimée en Bark^2 .

Pour notre étude, nous avons effectué un calcul de l'enveloppe convexe légèrement différent. En effet, la méthode que nous venons de présenter explique la nécessité d'utiliser tous les points (i.e., les voyelles) pour obtenir l'aire globale qui les enveloppe. Ce que nous cherchons à quantifier est l'espace vocalique d'un locuteur ou d'une langue en fonction des « timbres » vocaliques. Ainsi, chaque timbre vocalique est obtenu à partir des valeurs de F_1 & F_2 et ensuite on procède au calcul de l'enveloppe convexe. De cette manière, on obtient le plus petit polygone qui englobe tous ces timbres, et effectuons par la suite le calcul de l'aire de dispersion par la méthode proposée plus haut. Pour un exemple, voir la Figure 30, qui montre l'espace vocalique obtenu à partir de toutes les voyelles produites par le locuteur FRM03 dans la situation de production Mot en contexte /d/. Dans ce schéma, nous pouvons voir l'espace vocalique quadrilatère convexe obtenu pour toutes les voyelles (en ligne

continue) et l'enveloppe convexe des timbres vocaliques (en ligne discontinue). L'aire de dispersion des voyelles produites par ce locuteur et obtenue pour l'enveloppe convexe des timbres vocaliques est de 11,65 Bark².

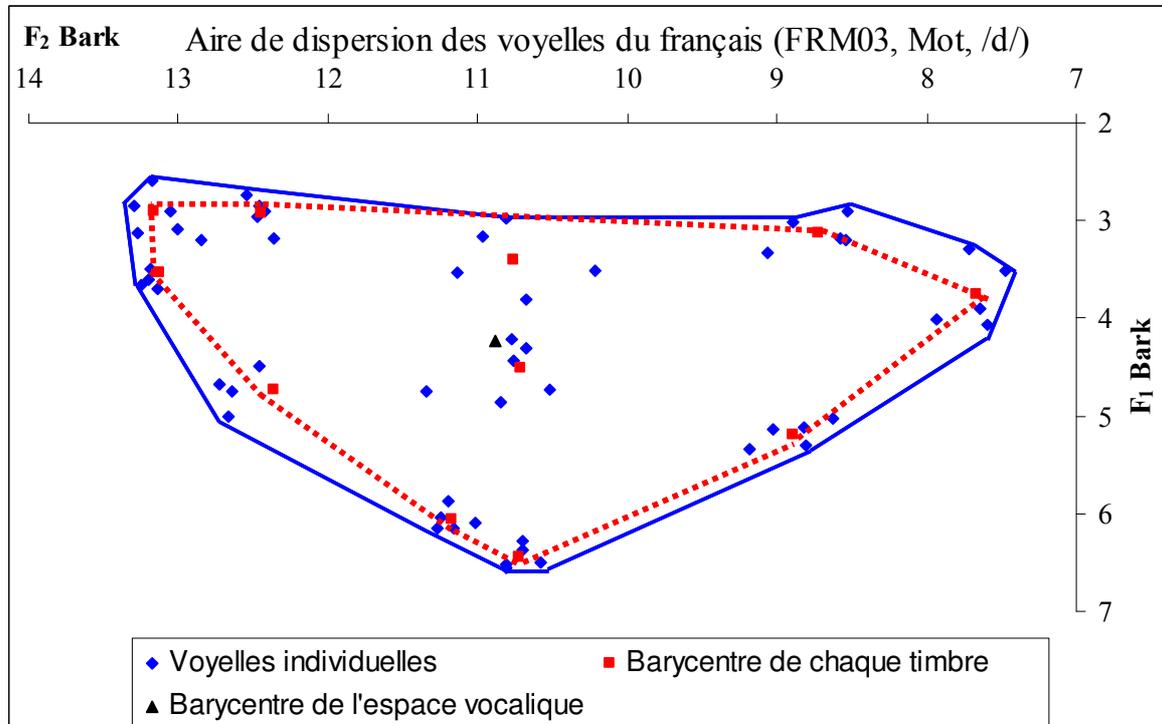


Figure 30 : Schéma représentant la dispersion des voyelles du FR produites par le locuteur FRM03 dans la situation de production Mot en contexte/d/ : en ligne bleue continue est représentée l'espace vocalique quadrilatère et en ligne rouge discontinus, l'Enveloppe Convexe de cet espace vocalique.

Une fois l'aire de dispersion obtenue, nous calculons le Barycentre de l'espace vocalique et les distances euclidiennes entre chaque voyelle et le Barycentre de l'espace correspondant. Le Barycentre correspond donc à la moyenne des valeurs formantiques des timbres vocaliques ; et la distance euclidienne, pour une voyelle par rapport à ce Barycentre, est obtenue à partir de la formule :

$$D_{\text{distance Euclidienne}} V_{\text{oyelle}} - B_{\text{arycentre}} = (F_1 V_{\text{oyelle}} - F_1 B_{\text{arycentre}})^2 + (F_2 V_{\text{oyelle}} - F_2 B_{\text{arycentre}})^2$$

Ces différents calculs permettaient de définir d'une part le « point central » ou le **centre de gravité** de l'espace vocalique, le Barycentre ; et d'autre part d'inclure et/ou d'exclure certaines voyelles s'éloignant trop du Barycentre et pouvant changer ainsi la taille de l'aire obtenue. Dans ce dernier cas, un nouveau calcul des Barycentres de chaque timbre vocalique et de l'enveloppe convexe est effectué.

Nous nous sommes inspiré du calcul de l'enveloppe convexe que nous avons appliqué sur les voyelles produites par langues et/ou locuteurs pour quantifier la taille des ellipses de dispersion de chaque catégorie vocalique par langue. Ainsi, nous effectuons le calcul du Barycentre pour les voyelles dans une catégorie précise pour chacun des locuteurs des trois langues dans un environnement et situation de production précis, et nous obtenons par conséquent dix Barycentres différents pour chaque catégorie vocalique. Ensuite nous calculons l'enveloppe convexe pour les dix timbres vocaliques et obtenons la taille de l'ellipse par la fonction « polyarea ». La taille de l'ellipse sera exprimée en Bark².

Afin de vérifier si les différences observées entre les paramètres testés sont significatives, nous avons appliqué une analyse statistique à nos données : une MANOVA à quatre facteurs (langue (3), situation de production (3), lieux d'articulation (4 en AM & AJ et 3 en FR) et voyelles (5 en AM, 8 en AJ et 11 en FR)) ainsi que toutes les interactions entre ces facteurs, effectuée à partir du logiciel NCSS© (Hintze, 2004).

Afin de tester la validité de nos résultats, nous avons appliqué une méthode de classification des données, une analyse discriminante, effectuée à partir du logiciel SPSS© (version 11.0.0., 2001). Cette analyse est une technique statistique qui vise à décrire, expliquer et prédire l'appartenance d'un ensemble d'observations (fréquences formantiques par exemple) à des groupes prédéfinis (langues, situations de production, voyelles, etc...) à partir de différentes variables prédictives classées en fonction de l'ordre d'utilisation (la durée par exemple) qui sont utilisées pour effectuer la classification. Dans notre étude, l'analyse discriminante est utilisée comme une méthode prédictive, car elle va permettre de construire une fonction de classification qui permet de prédire le groupe d'appartenance d'un individu (ou d'une voyelle par exemple) à partir des valeurs prises par les variables prédictives. Cette technique se rapproche des techniques supervisées en apprentissage automatique telles que les arbres de décision, les réseaux de neurones, etc...

Ainsi, ces deux analyses « statistiques » proposées permettent, d'un côté, de rendre compte des différences observées et de leurs validités statistiques, évaluées à partir de la MANOVA ; et de l'autre, de proposer de tester la pertinence de ces résultats en évaluant les distances entre les voyelles, par exemple, et de procéder à une classification de celles-ci.

Dans la partie qui suit, nous présentons les résultats de la comparaison entre les trois langues à partir des productions des trois voyelles /i a u/. Cette comparaison a pour but la mise en évidence des effets de la densité des systèmes vocaliques sur la taille des aires de dispersion ; sur la position des voyelles « extrêmes » et sur la taille des ellipses de chacune des trois voyelles. Comme nous le savons, l'espace vocalique dans chacune des trois langues n'est pas constitué uniquement des trois voyelles /i a u/. Ainsi nous vérifierons l'apport des autres voyelles dans l'agrandissement ou le rétrécissement des espaces vocaliques.

4.3.1.2. Résultats de dispersion des espaces vocaliques

Nous comparons dans cette partie la dispersion et la taille des espaces vocaliques des trois systèmes étudiés. Lorsque nous comparons les espaces vocaliques obtenus en fonction des langues et de la situation de production, nous observons des différences importantes corrélées, à notre sens, à la densité des systèmes. Les différents espaces vocaliques, présentés sur la Figure 31, montrent que globalement l'espace vocalique en FR est plus grand que celui de l'AM. Nous obtenons une différence significative en fonction de la taille de l'aire moyenne entre le FR et l'AM, $F(1, 18) = 7,43$; $p < 0,02$ (avec une moyenne d'aire de 7,11 Bark² en AM ; 8,14 Bark² en AJ et 8,95 Bark² en FR). La position des voyelles entre les trois langues est significativement différente en fonction de la situation de production. Ainsi sur F_1 , les voyelles fermées en FR sont plus fermées et les ouvertes, plus ouvertes ($F(4) = 6,67$; $p < 0,001$), tandis que sur F_2 et dans les trois situations de production, les voyelles postérieures en FR sont plus antérieures et les antérieures, moins antérieures ($F(4) = 58,99$; $p < 0,001$). L'effet de l'espace vocalique plus grand en FR par rapport aux deux autres langues est observé uniquement sur l'axe F_1 ; sur F_2 , c'est l'effet inverse.

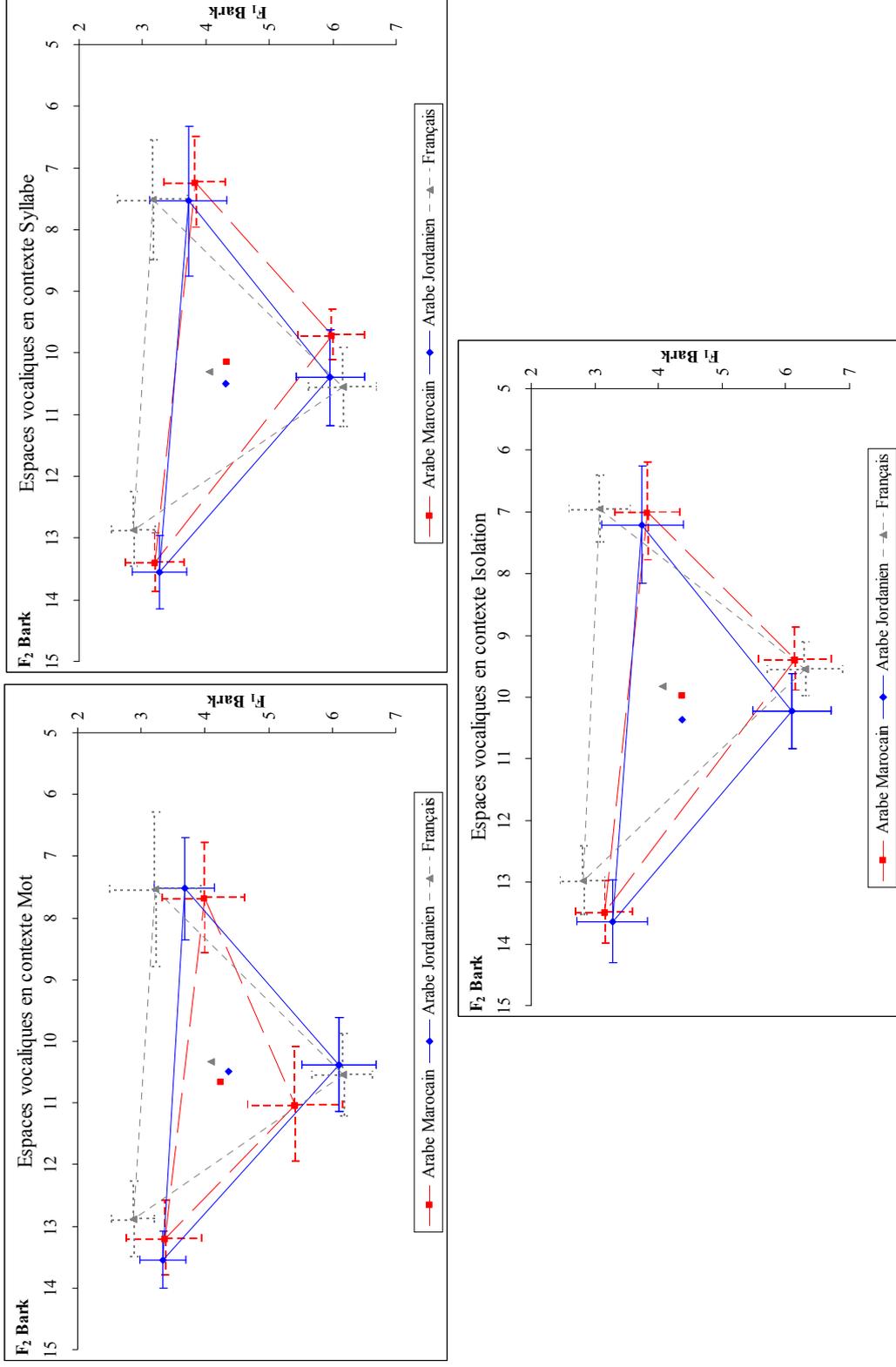


Figure 31 : Dispersion des trois timbres /i a u/ en AM, AJ et FR en fonction de la situation de production (voyelles produites dans les environnements /b d k/). Les points au centre des espaces vocaliques représentent les Barycentres pour chaque langue.

Nous observons que le Barycentre, ou **centre de gravité**, obtenu pour chaque langue et dans les trois situations de production est significativement différent ($p < 0,001$). Nous pouvons corrélé la position du Barycentre au déplacement dans les deux axes, ainsi, un Barycentre plus fermé indique un espace vocalique plus « fermé », et un Barycentre plus antérieur, indique un espace plus « antériorisé ». Les résultats obtenus ont montré que l'agrandissement de l'espace vocalique observé en FR est visible uniquement sur F_1 , les Barycentres obtenus montrent le même effet.

Bien que l'espace vocalique plus grand en FR soit obtenu uniquement sur l'axe F_1 ; la taille des aires de dispersion moyenne⁴⁸, dans les trois langues en fonction des situations de production, est significativement différente pour certaines situations de production (voir Figure 32, partie a). En situation de production Mot, l'aire de dispersion est significativement plus grande en AJ par rapport à celle en AM ($F(1, 18) = 11,28$; $p < 0,005$) ; en FR, elle est significativement plus grande que celle en AM ($F(1, 18) = 15,46$; $p < 0,002$) ; aucune différence entre les aires en AJ et FR ($F(1, 18) = 0,13$; $p = 0,72$). En Syllabe, aucune différence significative n'est obtenue entre les aires de dispersion dans les trois langues. En Isolation enfin, l'aire de dispersion est significativement plus grande en FR par rapport à l'AM ($F(1, 18) = 4,63$; $p < 0,05$) ; aucune différence entre les aires en AJ et AM ou les aires en AJ et FR. Lorsque les voyelles /i a u/ produites dans l'environnement /d^h/ sont ajoutées, aucun changement de taille de l'aire de dispersion n'est observé entre l'AJ et l'AM sauf dans la situation Syllabe, où nous obtenons une aire de dispersion moyenne non significativement plus petite en AM, alors qu'elle était plus grande avant, par rapport à l'AJ, ($F(1, 18) = 0,02$; $p = 0,9$), voir Figure 32, partie b). En conclusion, l'aire de l'espace vocalique en FR est significativement plus grande que celle dans les deux autres langues seulement en Mot.

⁴⁸ Nous utiliserons désormais les termes : **aire de dispersion des voyelles** pour parler de la taille de l'espace vocalique obtenue à partir des voyelles étudiées ; et **ellipse de dispersion des voyelles** pour parler de la taille des ellipses de chaque voyelle.

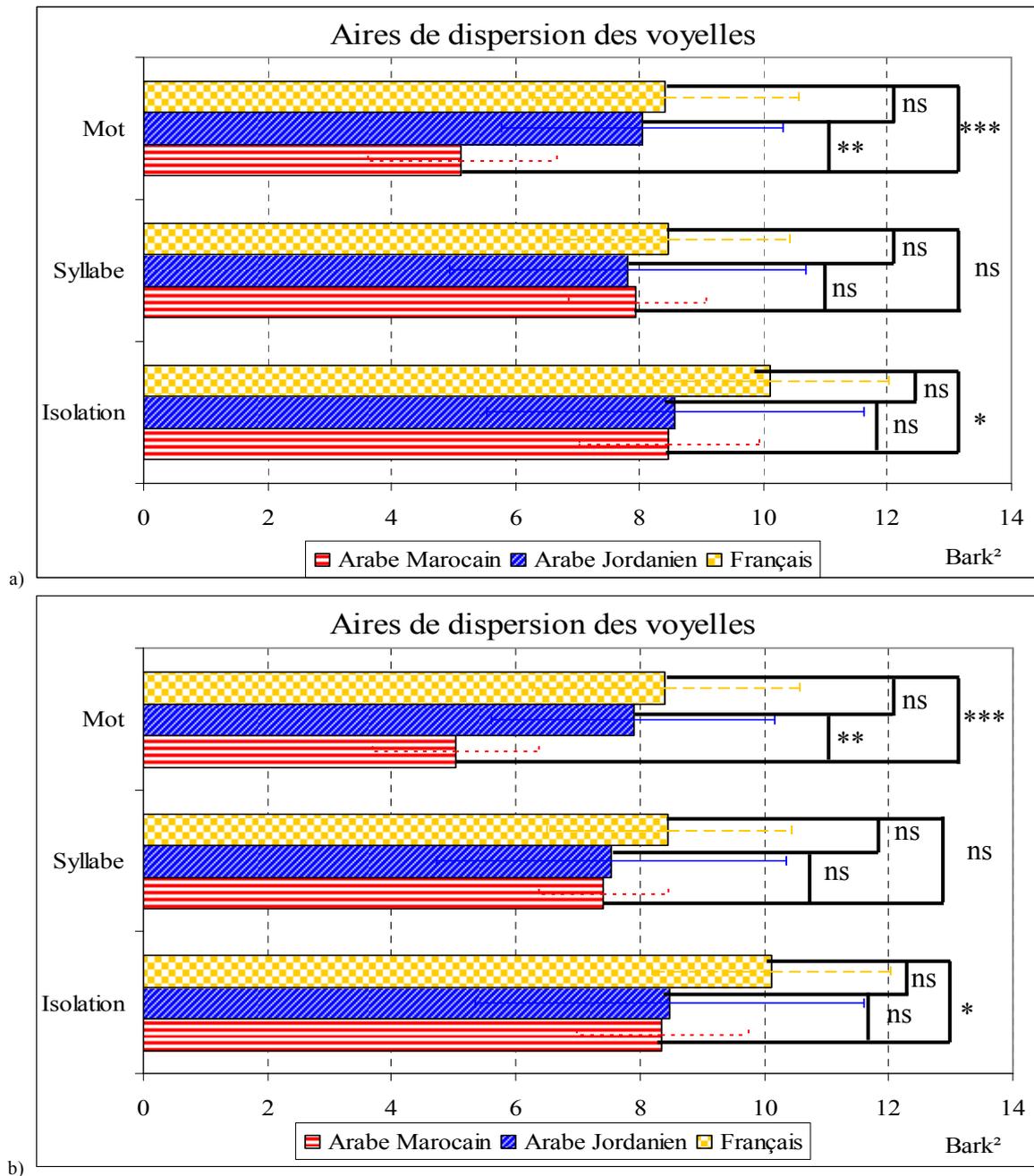


Figure 32 : Aires de dispersion (moyennes et écart-types) obtenues pour les voyelles /i a u/ en AM, AJ et FR en fonction de la situation de production, dans les environnements /b d k/ (partie a) et /b d d^s k/ (partie b). (ns = non significatif ; * = p<0,05 ; ** = p<0,01 ; *** = p<0,001).

Les écarts-types obtenus sont très grands, indiquant une variabilité inter-individuelle élevée au niveau de la réalisation des voyelles. Cette variabilité est très marquée en AJ (voir Tableau 6), où nous observons une augmentation de près de 300% de la taille de l'aire de dispersion en contexte Syllabe entre le locuteur 10 et le locuteur 6, la différence la moins marquée est observée en AM avec une différence de 28% entre les locuteurs 4 et 5.

Tableau 6 : Aires de dispersion des voyelles /i a u/ (en Bark²) en AM, AJ et FR en fonction de la situation de production, avec la taille de l'aire la plus petite et la plus grande et le numéro de locuteur (entre parenthèse).

	AM		AJ		FR	
	aire minimale	aire maximale	aire minimale	aire maximale	aire minimale	aire maximale
Mot	3,14 (10)	7,70 (9)	4,13 (5)	10,79 (6)	4,89 (7)	10,91 (6)
Syllabe	6,47 (8)	9,40 (1)	2,90 (10)	11,62 (9)	5,65 (10)	11,12 (8)
Isolation	6,22 (1)	10,34 (9)	3,89 (5)	11,97 (6)	6,12 (10)	12,6 (3)

Les aires de dispersion de l'espace vocalique des trois voyelles /i a u/ dans les trois langues ne sont pas très différentes : nous avons obtenu un espace vocalique plus grand uniquement en FR par rapport à l'AM, aucun effet significatif n'a été observé entre l'AM et l'AJ (sauf en situation Mot) ; les différences entre les aires de l'AJ et du FR ne sont pas significatives (sauf en situation Isolation). Afin de mesurer si ces effets sont identiques (ou pas) lorsque les autres voyelles sont examinées, nous avons comparé la taille des aires de dispersion de toutes les voyelles (/i a u/ incluses) entre les trois langues et quel que soit la situation de production (pour visualiser les différences des espaces vocaliques dans les trois langues, voir Annexe 1). Les résultats graphiques présentés par la Figure 33 (partie a) montrent que l'aire de dispersion moyenne en FR, dans les trois situations de production, est significativement plus grande que celle en AM ou en AJ. En situation de production Mot, l'aire de dispersion est significativement plus grande en AJ qu'en AM ($F(1, 18) = 4,61$; $p < 0,05$) ; en FR, elle est significativement plus grande qu'en AM ($F(1, 18) = 14,36$; $p < 0,002$) et aucune différence entre les aires en AJ et FR n'a été obtenue. En Syllabe, l'aire est significativement plus grande en FR par rapport à celle en AM uniquement ($F(1, 18) = 5,69$; $p < 0,05$) ; aucune différence entre les aires obtenues en AJ et en AM ou entre l'AJ et le FR. En Isolation enfin, l'aire est significativement plus grande en FR par rapport à l'AM ($F(1, 18) = 20,00$; $p < 0,001$) ; elle est significativement plus grande en FR par rapport à l'AJ ($F(1, 18) = 7,81$; $p < 0,05$) et enfin, aucune différence entre les aires en AJ et AM n'a été obtenue. Lorsque les voyelles en AJ et AM produites dans l'environnement /d^ɛ/ sont ajoutées (voir Figure 33, partie b), nous obtenons toujours les mêmes effets en FR (i.e., aire significativement plus grande que celle en AM ou AJ). Une différence significative entre les aires en AM et AJ dans la situation de production Syllabe est observée : l'aire est significativement plus grande en AJ ($F(1, 18) = 8,78$; $p < 0,01$). L'ajout de toutes les voyelles

a complètement modifié les résultats. Ainsi, nous obtenons une aire de dispersion significativement plus grande en FR par rapport aux deux autres langues.

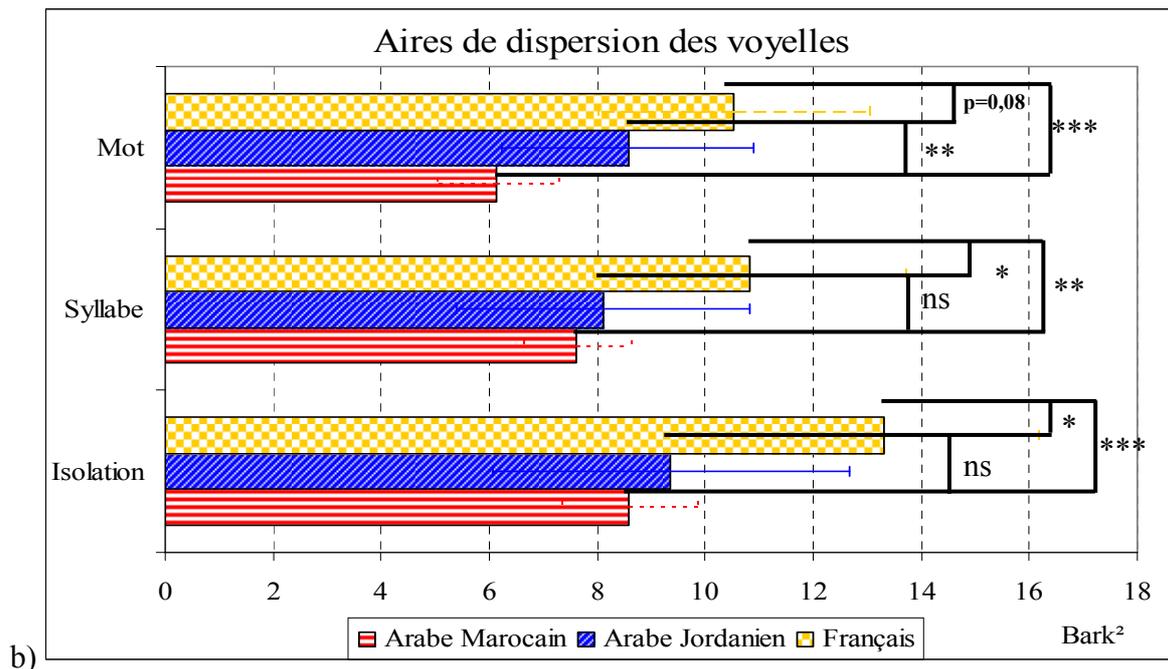
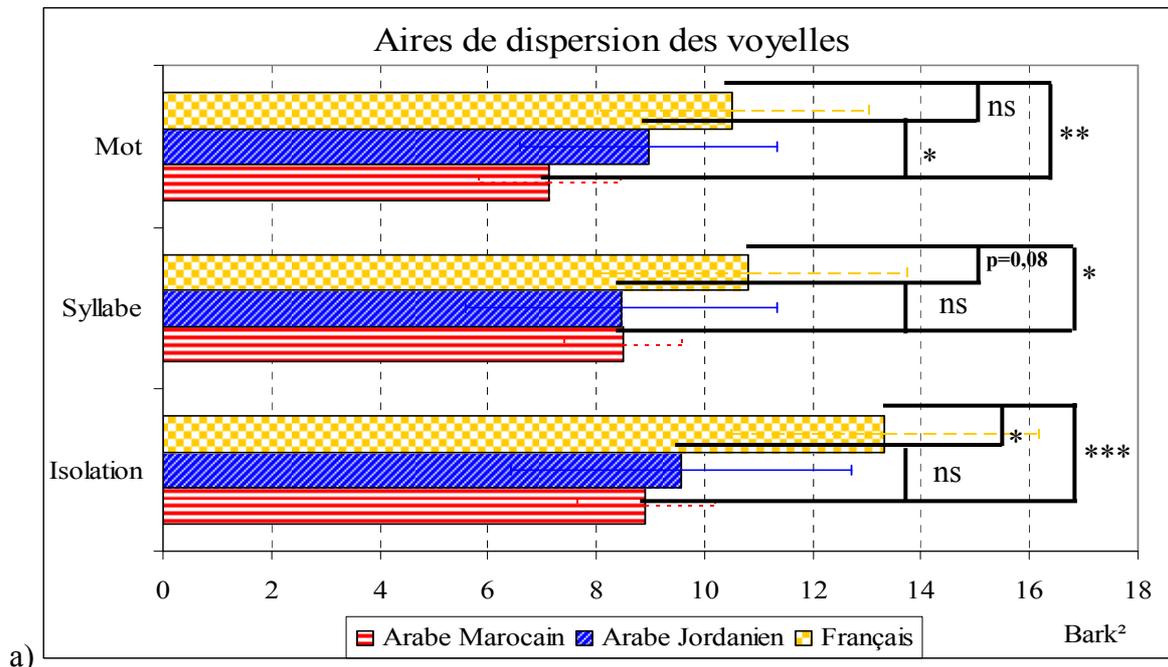


Figure 33 : Aires de dispersion (moyennes et écart-types) obtenues pour toutes les voyelles en AM, AJ et FR en fonction de la situation de production, dans les environnements /b d k/ (sous a) et /b d d^x k/ (sous b). (ns = non significatif ; * = p<0,05 ; ** = p<0,01 ; *** = p<0,001)

Nous obtenons approximativement les mêmes effets lorsque nous nous intéressons aux espaces vocaliques en fonction de la situation de production et du lieu d'articulation (voir Figures 34 à 36). Les espaces vocaliques sont significativement plus grands en FR et plus

petits en AM ($p < 0,001$). L'espace vocalique plus grand en FR est obtenu en fonction de la position des trois voyelles /i a u/ quels que soient la situation de production et le lieu d'articulation.

En ce qui concerne les différences liées aux lieux d'articulation, les résultats montrent que c'est le contexte /k/ dans les trois langues qui affecte le plus l'espace vocalique sur F_1 et ce quelle que soit la situation de production : les voyelles ouvertes sont légèrement moins ouvertes et les fermées, légèrement moins fermées ($F(12) = 8,59$; $p < 0,001$), tandis que sur F_2 , c'est le contexte /d/ dans les trois langues qui affecte le plus l'espace vocalique : les voyelles antérieures sont plus postérieures et les postérieures moins postérieures ($F(12) = 12,31$; $p < 0,001$).

Pour ce qui est des effets obtenus quels que soient la situation de production et le lieu d'articulation, c'est en situation Mot que les voyelles sont les plus affectées. Plus particulièrement, c'est en FR que les voyelles sont plus « périphériques » sur l'axe F_1 uniquement : /i u/ sont plus fermées et /a/ est plus ouverte en FR par rapport aux voyelles en AM ou en AJ. Nous avons examiné plus en détail ces différences, et nous présentons dans le Tableau 7 (p. 141) les résultats des tests statistiques sur les valeurs de F_1 & F_2 séparément, pour chaque lieu d'articulation en situation de production Mot.

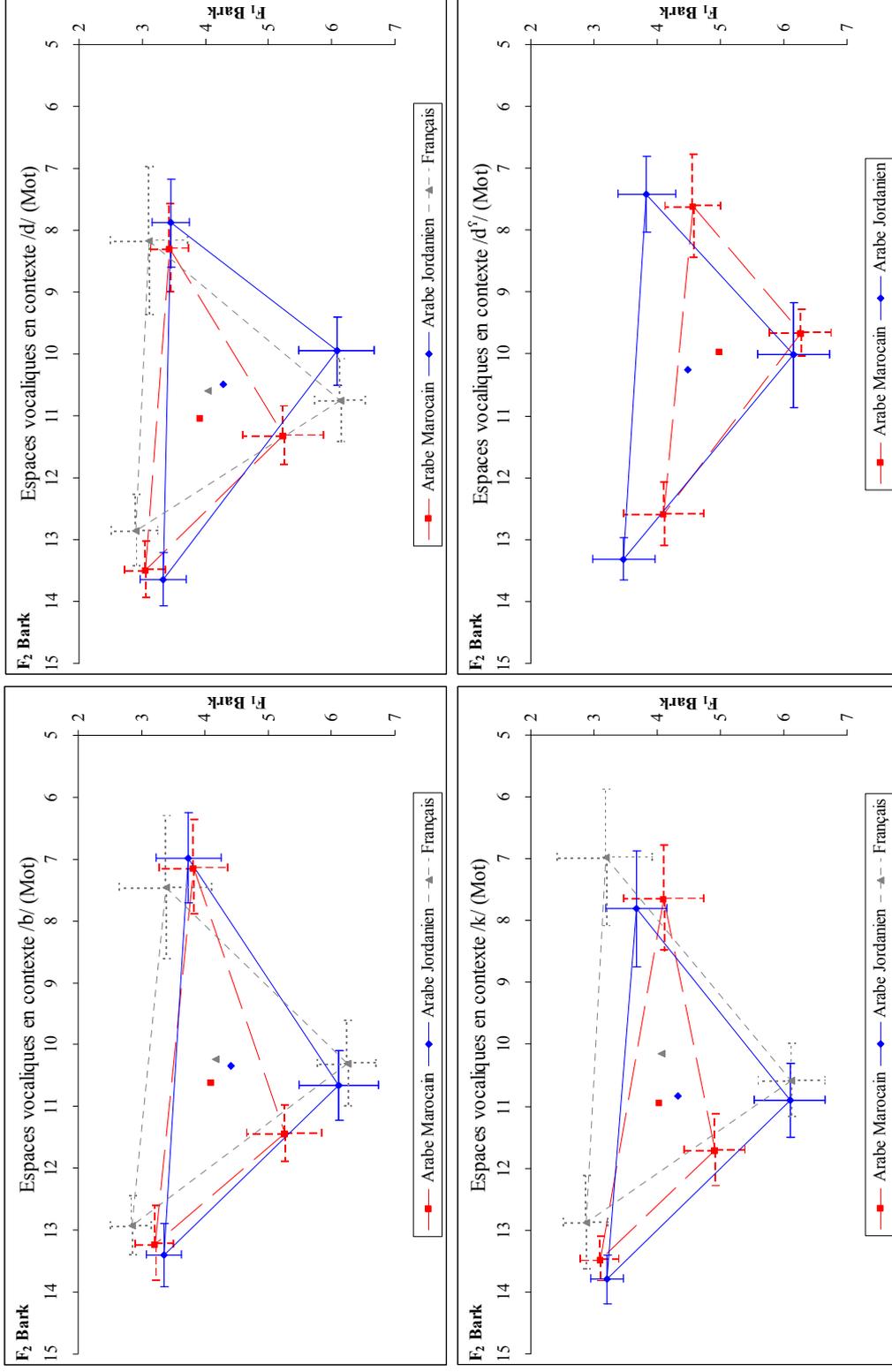


Figure 34 : Dispersion des trois timbres /i a u/ en AM, AJ et FR en fonction de la situation de production Mot et du lieu d'articulation. Les points au milieu de chaque espace vocalique représentent le Barycentre de l'espace pour chaque langue.

RÉSULTATS EN PRODUCTION DES VOYELLES

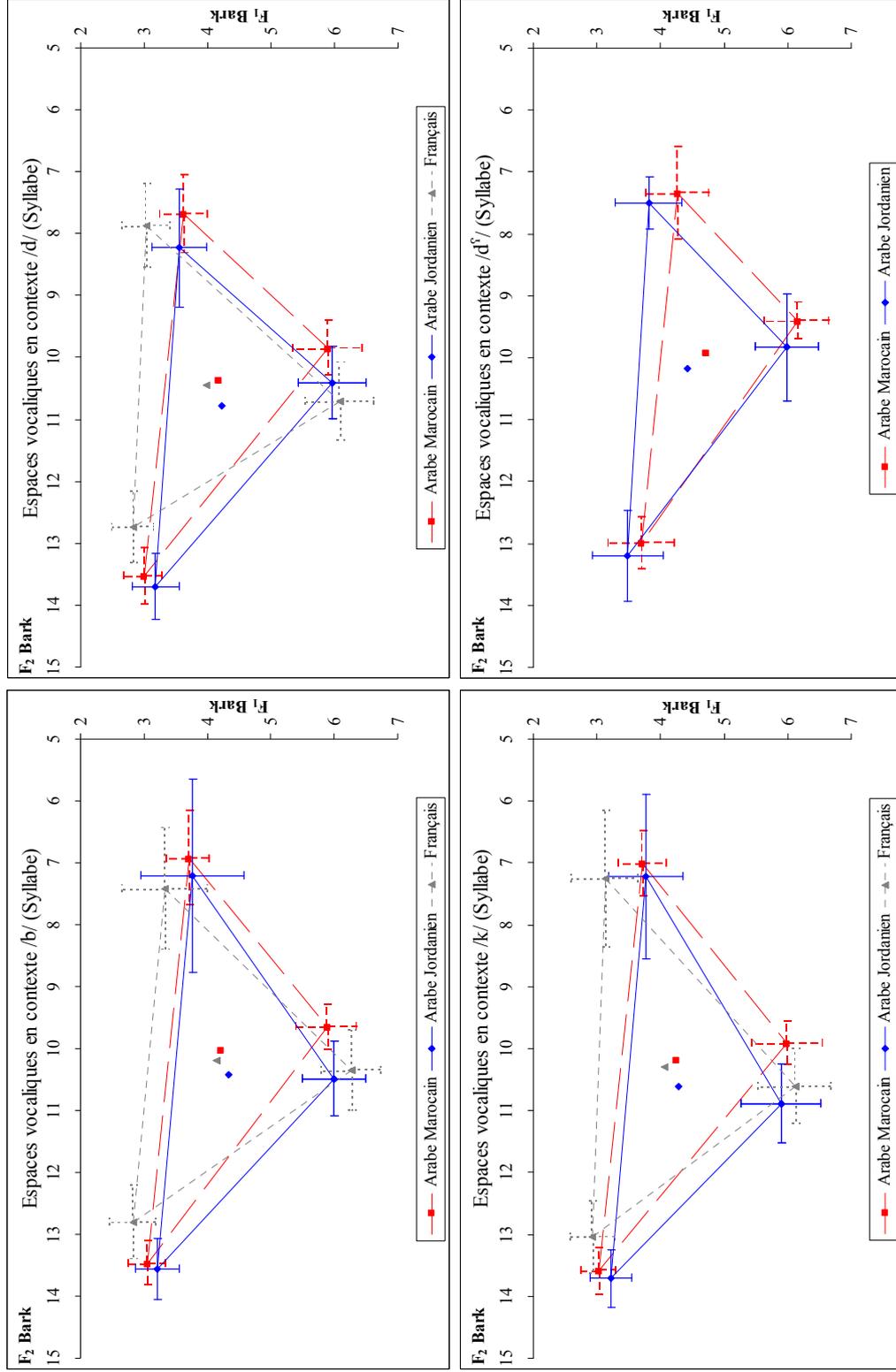


Figure 35 : Dispersion des trois timbres /i a u/ en AM, AJ et FR en fonction de la situation de production Syllabe et du lieu d'articulation. Les points au milieu de chaque espace vocalique représentent le Barycentre de l'espace pour chaque langue.

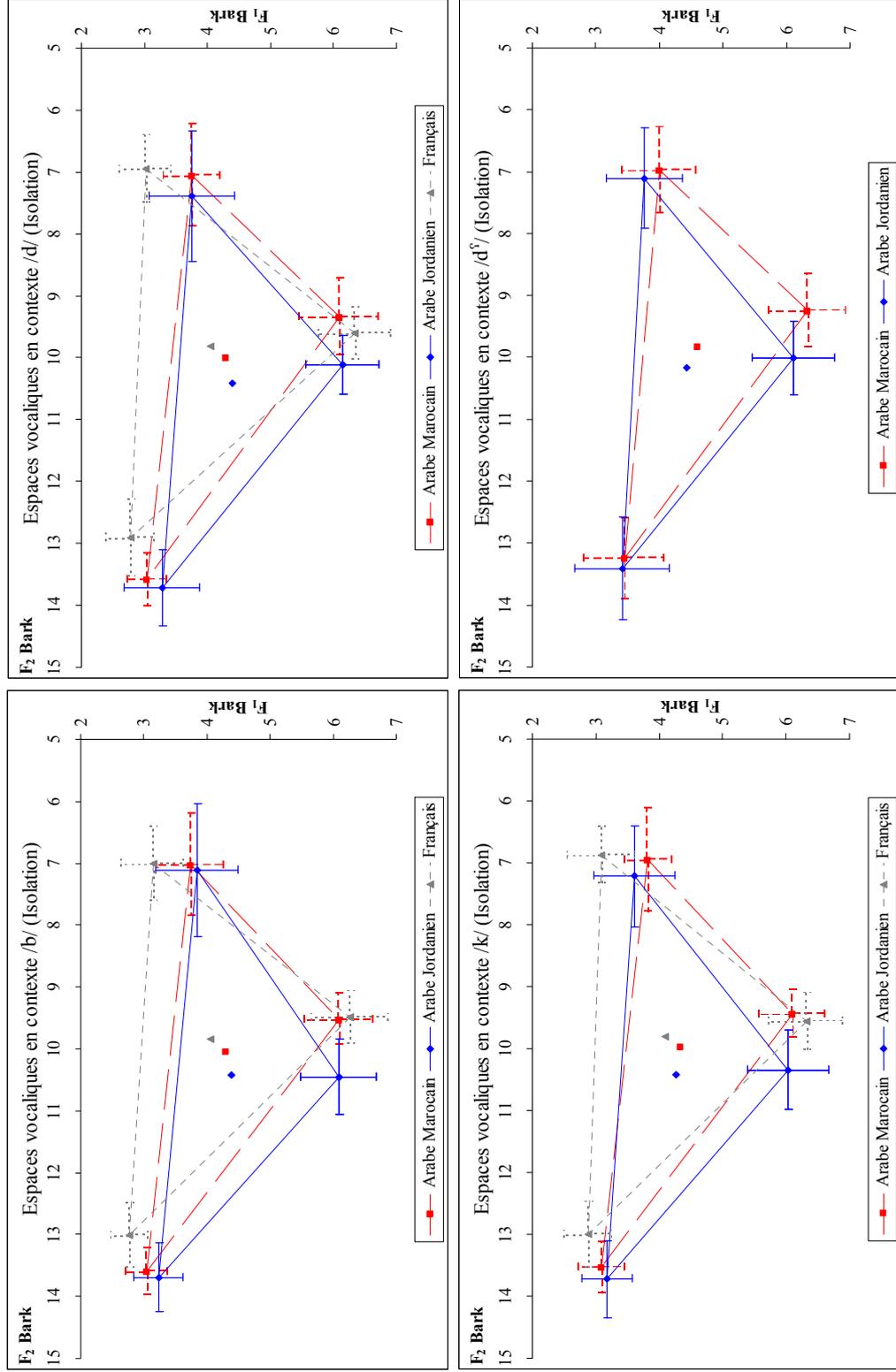


Figure 36 : Dispersion des trois timbres /i u/ en AM, AJ et FR en fonction de la situation de production Isolation et du lieu d'articulation. Les points au milieu de chaque espace vocale représentent le Barycentre de l'espace pour chaque langue.

Tableau 7 : Résultats des tests MANOVA entre les trois voyelles /i a u/ en AM, AJ et FR, en contexte Mot. [+F] = [+fermé], [+O] = [+ouvert], [+A] = [+antérieur], [+P] = [+postérieur], NS = non significative. Le tableau se lit de la façon suivante : /b/ > /i/ > AJ vs. AM = [+F], p<0,001 : la voyelle /i/ en contexte /b/ est [+fermée] en AJ par rapport à AM.

		F ₁			F ₂		
		AJ vs AM	AM vs FR	AJ vs FR	AJ vs AM	AM vs FR	AJ vs FR
/b/	/i/	[+F], p<0,001	[+O], p<0,001	[+O], p<0,001	[+A], p<0,01	[+A], p<0,001	[+A], p<0,001
	/a/	[+O], p<0,001	[+F], p<0,001	NS	[+P], p<0,001	[+A], p<0,001	[+A], p<0,001
	/u/	NS	[+O], p<0,001	[+O], p<0,001	NS	[+P], p<0,01	[+P], p<0,001
/d/	/i/	[+O], p<0,001	[+O], p<0,001	[+O], p<0,001	[+A], p<0,01	[+A], p<0,001	[+A], p<0,001
	/a/	[+O], p<0,001	[+F], p<0,001	NS	[+P], p<0,001	[+A], p<0,001	[+P], p<0,001
	/u/	NS	[+O], p<0,001	[+O], p<0,001	[+P], p<0,001	NS	NS
/k/	/i/	[+O], p<0,01	[+F], p<0,001	[+O], p<0,001	[+A], p<0,001	[+A], p<0,001	[+A], p<0,001
	/a/	[+O], p<0,001	[+F], p<0,001	NS	[+P], p<0,001	[+A], p<0,001	[+A], p<0,001
	/u/	[+F], p<0,001	[+O], p<0,001	[+O], p<0,001	[+A], p<0,001	[+A], p<0,001	[+A], p<0,001
/d ^s /	/i/	[+O], p<0,001			[+A], p<0,001		
	/a/	NS			[+A], p<0,001		
	/u/	[+O], p<0,001			NS		

Les aires de dispersion des voyelles /i a u/ obtenues en fonction de la situation de production et du lieu d'articulation sont présentées sur la Figure 37. Les trois graphiques présentés sous a, b & c, montrent les aires obtenues à partir des voyelles produites dans les environnements /b d k/. Dans la situation de production Mot et en contexte /b/, l'aire est significativement plus petite en AM par rapport à celle en AJ ($F(1, 18) = 8,09$; $p<0,02$) ou celle en FR ($F(1, 18) = 11,01$; $p<0,005$) ; en contexte /d/, elle est significativement plus petite en AM par rapport à celle en AJ ($F(1, 18) = 8,59$; $p<0,01$) ou celle en FR ($F(1, 18) = 4,99$; $p<0,05$) ; en contexte /k/, elle est significativement plus petite en AM par rapport à celle en AJ ($F(1, 18) = 14,88$; $p<0,002$) ou celle en FR ($F(1, 18) = 30,16$; $p<0,001$). Dans la situation de production Syllabe quel que soit les trois lieux d'articulation, aucune différence significative entre les aires de dispersion n'est obtenue. Autrement dit, nous obtenons la même dispersion quelle que soit la langue. Enfin, dans la situation de production Isolation, l'aire de dispersion des voyelles /i a u/ en AM est significativement plus petite par rapport à celle en FR dans l'environnement /k/ ($F(1, 18) = 4,64$; $p<0,05$) ; une tendance est obtenue en contexte /d/ : l'aire est plus petite en AM par rapport à celle en FR ($F(1, 18) = 4,30$; $p=0,053$). Nous n'obtenons aucune différence significative entre les aires de dispersion des trois

voyelles /i a u/ en AJ ou en FR, quels que soient la situation de production et le lieu d'articulation. Le graphique présenté sur la Figure 37 (partie d) montre les aires de dispersion des voyelles en AM et en AJ dans l'environnement /d^h/. L'aire de dispersion des voyelles dans cet environnement et dans la situation de production Mot est significativement plus grande en AJ par rapport à celle en AM ($F(1, 18) = 8,40; p < 0,01$) ; dans les deux situations de production Syllabe et Isolation, la différence n'est pas significative.

Ces résultats montrent que globalement, les trois voyelles « extrêmes » produites, dans les trois situations de production en fonction des environnements consonantiques, sont réalisées différemment dans les trois langues. Plus particulièrement, c'est dans la situation Mot que les effets de la densité des systèmes sont les plus visibles, tant sur la position de ces voyelles, que par rapport à la taille de l'aire. Dans les situations Syllabe et Isolation, la position des voyelles et la taille de l'aire de dispersion semblent être « identiques » dans les trois langues quel que soit le lieu d'articulation. Ainsi les espaces vocaliques obtenus pour les trois voyelles extrêmes, dans ces deux situations de production, ne montrent pas les effets de la densité des systèmes.

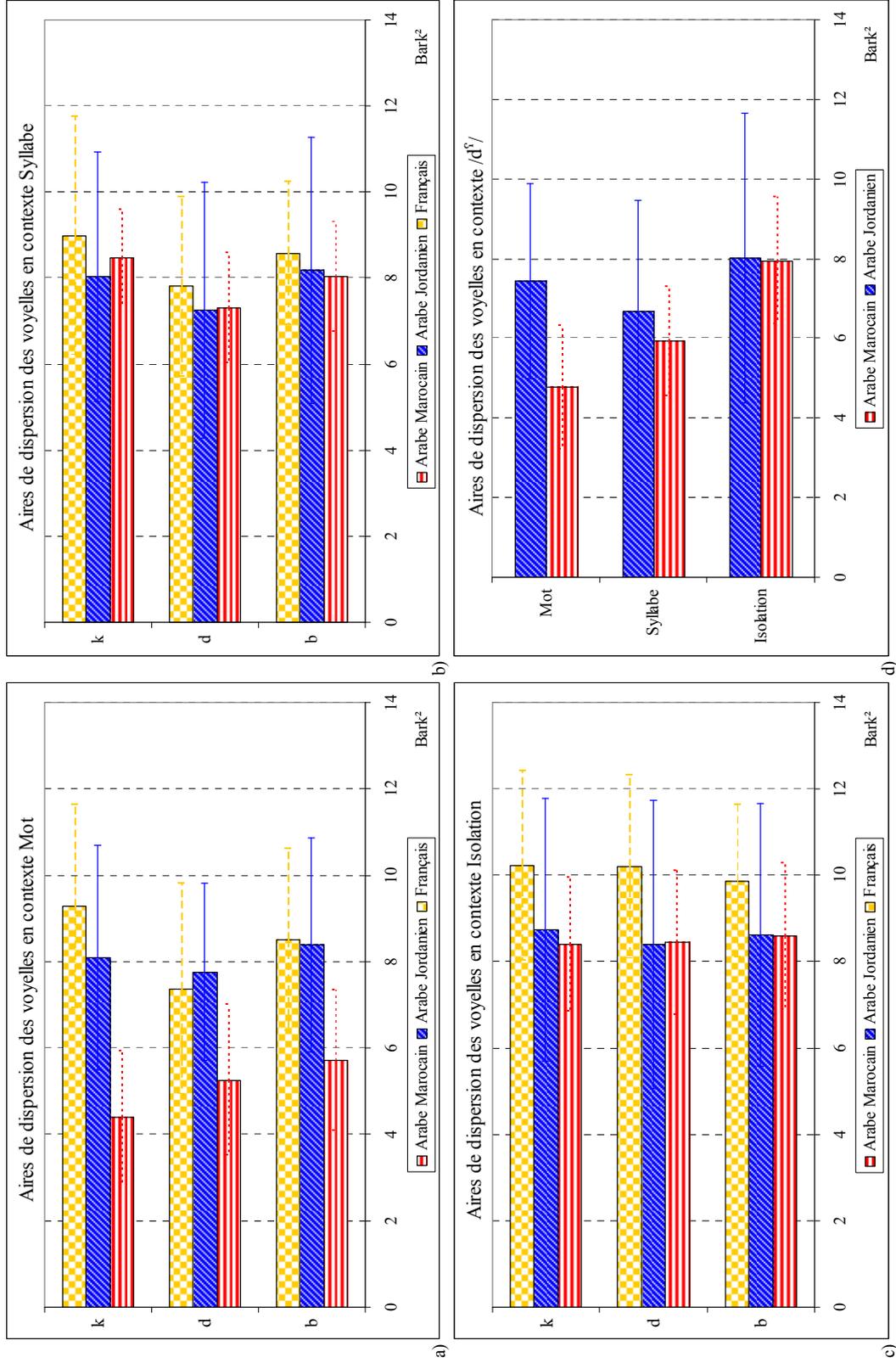


Figure 37 : Aires de dispersion des voyelles en AM, AJ et FR en fonction de la situation de production et du lieu d'articulation de /b d k/. Les aires en contexte /d^h/ en AM et AJ sont présentées sous (d).

Afin de mieux visualiser l'effet de la taille des espaces vocaliques en FR ou en AJ par rapport à celle en AM, nous présentons, dans le Tableau 8, les différences relatives observées. En considérant la taille de l'espace vocalique en AM comme correspondant à 100%, nous la comparons à la taille de l'espace vocalique en AJ ou en FR. Nous observons une expansion en moyenne de 158,50% en Mot, de 102,10% en Syllabe et de 101,25% en Isolation de l'espace vocalique en AJ par rapport à celui en AM, et de 166,74% en Mot, de 106,52% en Syllabe et de 119,24% en Isolation de l'espace vocalique en FR par rapport à celui en AM. La comparaison avec AJ montre que l'espace vocalique en FR est plus grand en moyenne de 103,77% en Mot, de 108,11% en Syllabe et de 117,65% en Isolation. Il est à noter que les aires de dispersion des voyelles /i a u/, dans les trois situations de production et quel que soit le lieu d'articulation, sont toujours plus grandes en FR par rapport à celles en AM ($FR > AM$) ; que dans certains contextes, les aires obtenues en AJ et en FR sont comparables ($FR \approx AJ$) tandis que les aires en AJ sont quelques fois plus petites que celles en AM ($AJ < AM$).

Tableau 8 : Correspondance en pourcentage de taille de l'aire de dispersion des voyelles /i a u/ en AJ et en FR en comparaison avec celle en AM, considérée à 100%. Les flèches indiquent lorsque le pourcentage est supérieur en AJ ou en FR par rapport à celui en AM (flèche montante) ; lorsqu'il est inférieur (flèche descendante) ou lorsqu'il est « presque » identique (flèche droite).

		Arabe Marocain	Arabe Jordanien	Français
Mot	b	100,00%	146,81% ↗	148,96% ↗
	d	100,00%	147,56% ↗	140,23% ↗
	k	100,00%	183,79% ↗	211,02% ↗
	d^s	100,00%	155,85% ↗	
Syllabe	b	100,00%	101,77% →	106,81% ↗
	d	100,00%	99,20% ↘	106,82% ↗
	k	100,00%	94,83% ↘	105,93% ↗
	d^s	100,00%	112,60% ↗	
Isolation	b	100,00%	100,45% →	114,97% ↗
	d	100,00%	99,44% ↘	120,93% ↗
	k	100,00%	104,21% ↗	121,81% ↗
	d^s	100,00%	100,89% →	

Les comparaisons effectuées au niveau de la dispersion des voyelles, de la taille des aires de dispersion et des effets d'expansion de l'espace vocalique en AJ et en FR par rapport

à l'espace vocalique en AM vont globalement dans le sens d'une expansion de la taille de l'espace vocalique est fonction de la densité des systèmes vocaliques, et ainsi dans le sens de la **théorie de la dispersion adaptée**, comme proposée par Lindblom (1986). Nous observons également des effets significatifs de la situation de production compatible avec une analyse en termes de quantité d'informations contextuelle véhiculée : plus la quantité d'informations contextuelle véhiculée est importante (comme dans la réalisation des voyelles en Mot), plus la variabilité est importante et plus la dispersion des espaces vocaliques est grande par rapport à la réalisation des mêmes voyelles en Syllabe ou en Isolation. Ces résultats vont dans le sens des propositions de Lindblom (1990) dans sa théorie *H&H* : la réalisation des voyelles en situation Mot correspond à une forme *Hypo-Articulée* des voyelles ; celle en situation Isolation correspond à une forme *Hyper-Articulée* des voyelles. Les effets du lieu d'articulation et/ou de l'interaction entre le lieu d'articulation et la situation de production ne sont pas prédits dans cette théorie. Néanmoins, nous n'observons pas une homogénéité entre nos résultats et ceux obtenus dans Stevens & House (1963*) où un effet de « centralisation » de l'espace vocalique est obtenu dans l'environnement /d/.

D'après les explications de la littérature, les effets de la densité des systèmes sur la dispersion des espaces vocaliques ne se limitent pas uniquement aux différences obtenues en fonction de la taille des espaces vocaliques et des positions relatives des timbres ; un effet d'agrandissement et/ou rétrécissement de la taille des ellipses de chaque catégorie vocalique est observé. Ainsi, nous nous intéressons dans la partie qui suit à cette question.

4.3.1.3. Résultats des ellipses de dispersion de chaque catégorie vocalique

Les influences de la densité des inventaires vocaliques ne se traduisent pas uniquement sur la taille et la dispersion des espaces vocaliques ; nous nous attendons à observer des différences liées à la taille des ellipses de dispersion de chaque catégorie vocalique. Selon la **théorie quantique** et la **théorie de la dispersion-focalisation**, les trois voyelles /i a u/, qui forment des « *Hot-Spots* » sont présentes dans des positions extrêmes de l'espace vocalique et présentent une dispersion intra-catégorie très faible. Selon la **théorie de la dispersion**

adaptée, ces trois voyelles doivent être distinctes les unes des autres, mais dépendent de la densité des systèmes vocaliques : plus une langue a de voyelles, plus l'espace vocalique est grand. Pour cette dernière théorie, on obtiendra deux cas de figures : l'espace vocalique plus réduit dans une langue à faible densité présente soit une dispersion intra-catégorie faible pour ne pas les confondre, car l'espace vocalique obtenu est très « centralisé » ; soit une dispersion plus grande, car la « centralisation » de l'espace vocalique n'est pas obtenue et par conséquent les voyelles « extrêmes » limitent l'espace vocalique. Nous examinons donc ici si la densité des systèmes vocaliques affecte la dispersion de chaque catégorie en AM, AJ et FR à partir des trois voyelles /i a u/.

Nous avons ainsi comparé les ellipses de dispersion dans les trois langues en fonction de la situation de production, (voir Figure 38). Globalement, la taille des ellipses de dispersion des catégories vocaliques en AM est la plus petite par rapport à celle en AJ ou en FR (sauf pour /u/ en Isolation). Ailleurs, les ellipses de dispersion des catégories vocaliques sont plus grandes par rapport à celles en FR (sauf pour /i u/ en Mot). Ce premier résultat montre que la dispersion intra-catégorie des trois voyelles /i a u/ semble ne pas dépendre de la densité des systèmes car nous obtenons effet inverse de celui attendu : nous nous attendions à obtenir des ellipses plus grandes en AM et plus petites en FR, or nous obtenons les ellipses les plus petites en AM.

Nous avons comparé la taille des ellipses de dispersion par situation de production et lieu d'articulation. Les résultats présentés dans le Tableau 9 montrent la même tendance : les ellipses de dispersion des voyelles /i a u/ en AM sont majoritairement plus petites que celles en AJ, le FR présente des ellipses, en moyenne, plus petites que celles obtenues pour l'AJ.

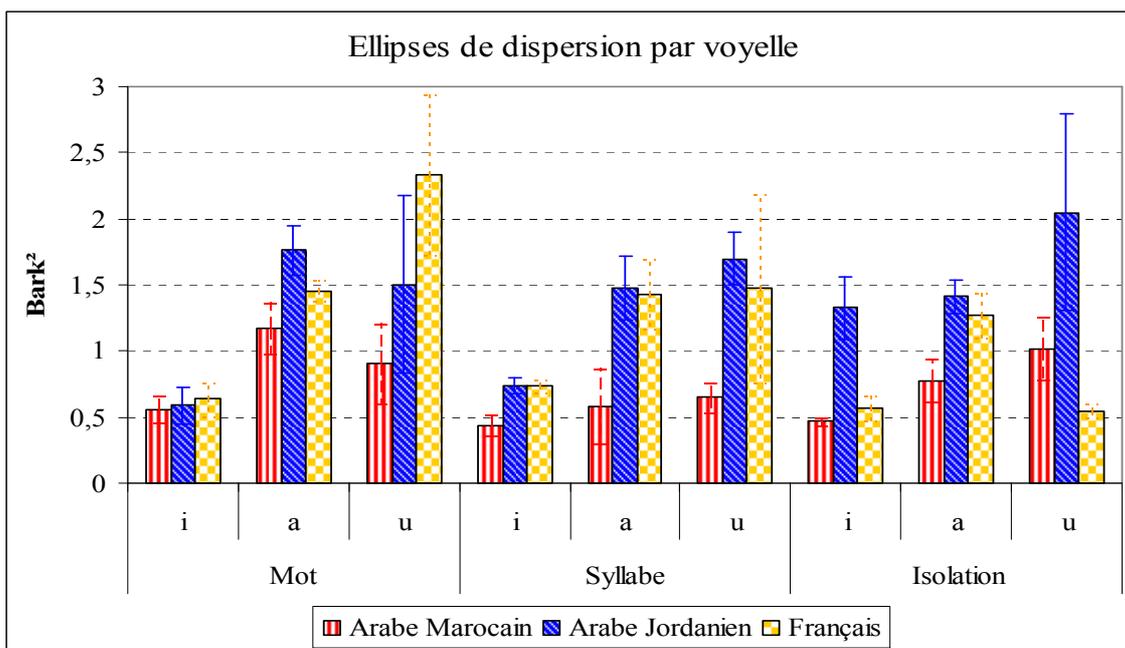


Figure 38 : Ellipses de dispersion des voyelles /i a u/ en AM, AJ et FR en fonction de la situation de production.

Tableau 9 : Taille des ellipses de dispersion des catégories vocaliques pour /i a u/ en AM, AJ et FR en fonction de la situation de production et du lieu d'articulation (en Bark²).

		i			a			u		
		AM	AJ	FR	AM	AJ	FR	AM	AJ	FR
mot	b	0,65	0,53	0,51	1,39	1,94	1,53	0,94	1,51	3,31
	d	0,58	0,75	0,75	1,11	1,79	1,37	0,59	0,83	1,77
	k	0,45	0,49	0,67	1,03	1,57	1,47	1,21	2,17	1,92
	d ^f	0,82	0,64		0,76	2,75		1,62	1,02	
Syllabe	b	0,46	0,81	0,79	0,52	1,45	1,24	0,52	1,84	1,25
	d	0,51	0,68	0,71	0,91	1,26	1,33	0,75	1,47	0,81
	k	0,36	0,74	0,72	0,34	1,74	1,73	0,67	1,79	2,38
	d ^f	0,81	1,16		0,47	2,48		1,72	0,79	
Isolation	b	0,51	1,19	0,47	0,82	1,56	1,08	1,16	2,59	0,52
	d	0,44	1,61	0,56	0,92	1,32	1,38	0,75	2,35	0,52
	k	0,45	1,19	0,66	0,61	1,35	1,35	1,16	1,21	0,61
	d ^f	0,85	2,74		0,87	1,64		1,37	1,28	

Ces résultats ne sont pas conformes aux propositions de la **théorie de la dispersion adaptée**, car même si nous obtenons des espaces vocaliques plus réduits en AM par rapport aux deux autres langues, la taille des ellipses de chaque voyelle semble ne pas dépendre de la densité des systèmes. En fonction des différents cas de figures que nous avons présentés sur la Figure 26 (p. 113), nos résultats vont dans le sens du Cas N°. 2 où la taille des espaces

vocaliques semble dépendre de la densité des systèmes avec une dispersion intra-catégorie vocalique faible.

L'observation que nous faisons des résultats des aires de dispersion des espaces et des ellipses de chaque catégorie vocalique de chaque système, semble montrer le rôle important de la variabilité inter-individuelle pour expliquer les résultats contradictoires obtenus. En effet, certains locuteurs obtiennent des aires de dispersion des espaces vocaliques très variables allant de 2 à 14 Bark² (suivant la langue, la situation de production et le lieu d'articulation), ainsi que des réalisations différentes des voyelles dans chaque catégorie pour les cinq répétitions, et présentant ainsi des écarts de fréquences de 1 Bark sur F₁ ou F₂. La prise en compte des résultats individuels devra permettre de mettre en avant les différentes stratégies employées par chaque locuteur dans ses réalisations vocaliques ; travail difficile à accomplir dans une étude comme la nôtre. Néanmoins, nous tâcherons de présenter quelques résultats individuels lorsque les spécificités de chaque langue seront abordées (voir § 4.4).

Nous avons proposé dans cette partie une analyse **statique** des formants des voyelles produites dans les trois langues et plus particulièrement les voyelles /i a u/ (voyelles longues en arabe) afin d'évaluer le rôle de la densité des systèmes sur la dispersion des voyelles, en termes de taille des espaces vocaliques, de positions relatives de ces voyelles et de taille d'ellipses de chaque catégorie vocalique. Globalement, ces résultats ont montré que le FR avait tendance à présenter des espaces vocaliques plus grands que ceux en AM ; l'espace vocalique en AJ se situe entre les deux. Cette analyse **statique** a permis de rendre compte de ces différences, mais comme nous l'avons expliqué dans le Chapitre 2, les voyelles sont mieux identifiées par les auditeurs à partir des indices **dynamiques** contenus dans le pattern spectral de celles-ci (soit à partir des Changements Spectraux Inhérents aux Voyelles (les CSIV), soit à partir des transitions et trajectoires des voyelles). Nous nous intéressons par conséquent dans la partie qui suit à l'analyse **dynamique** des voyelles en évaluant tout d'abord le rôle de la densité des systèmes sur les transitions formantiques.

4.3.2. *Analyses dynamiques des formants des voyelles*

Nous décrivons dans cette partie les systèmes vocaliques des trois langues en effectuant une analyse **dynamique**. Par analyse **dynamique** nous entendons une caractérisation **dynamique** des voyelles en prenant en compte plusieurs valeurs formantiques au cours de la réalisation des voyelles en comparaison avec l'utilisation d'une seule valeur correspondant généralement à la position centrale utilisée dans l'analyse **statique**. Plus particulièrement, nous nous intéressons à l'étude des effets de la densité des systèmes sur les transitions formantiques en étudiant les trois voyelles communes entre les trois langues /i a u/ pour permettre des comparaisons. Le but de cette analyse comparative est de mettre en avant le rôle des aspects **dynamiques** dans l'analyse des voyelles et les différences existantes entre les trois langues. L'évaluation de la pertinence de ces analyses **dynamiques** dans la description des voyelles /i a u/ dans les trois langues, en comparaison avec les valeurs **statiques** sera établie à partir des résultats de l'analyse discriminante.

4.3.2.1. **Analyse et traitement des données**

Nous avons analysé 11988 sur les 12036 voyelles étudiées précédemment et produites par les locuteurs des trois langues (soit 2907 voyelles en AM, 4283 en AJ et 4798 en FR). Pour pouvoir estimer correctement les paramètres **dynamiques**, nous avons imposé une durée minimale de voyelle de 40 ms, ce qui explique la différence du nombre de voyelles entre les deux analyses.

Les mesures acoustiques des trois premiers formants de chaque voyelle ont été effectuées avec le logiciel Praat en utilisant l'algorithme d'extraction de formants « Burg » (analyse LPC avec auto-corrélation, avec 24 coefficients) avec une fenêtre d'analyse gaussienne de 12,5 ms et un pas de déplacement de 5 ms. Les valeurs formantiques de l'**onset** vocalique ont été déterminées en suivant la méthode proposée dans Al-Tamimi (2004), et correspondent à la valeur distante de 5 ms du début de la transition vers la voyelle (i.e., la valeur obtenue à 5 ms après le début de la période vocalique). Les valeurs formantiques, extraites toutes les 5 ms, ont été vérifiées manuellement afin de corriger les possibles erreurs

d'extraction automatique. Ces valeurs ont été ensuite converties en Bark (par la même formule utilisée pour l'analyse **statique**). Nous avons choisi les aspects **dynamiques** de la production des voyelles par la **pente formantique** obtenue à partir d'une analyse de régression linéaire et polynomiale effectuée sur la transition débutant à l'**onset** pour se terminer à l'état stable. Ce calcul rend compte des influences consonantiques sur les voyelles, en modélisant les transitions de chaque formant.

Le calcul des **pentés formantiques** que nous nous proposons d'exploiter est dérivé de celui des **équations du locus** initialement proposé par Lindblom (1963_b), où des régressions linéaires dérivées de la relation entre les valeurs des **onsets** du deuxième formant de différentes voyelles (F_{2Onset}) et leurs valeurs de l'état stable correspondant ($F_{2Milieu}$) sont proposées. Pour pouvoir apprécier les différences entre les deux méthodes, nous allons les décrire plus en détails. Ainsi nous commençons par expliquer ce que nous entendons par « transition formantique », quelles sont les différences entre le « locus virtuel » (Couper *et al.*, 1952 ; Delattre *et al.*, 1952, 1955 ; entre autres) et les **équation du locus** (Lindblom, 1963_b, entre autres) ; et enfin comment les **pentés formantiques** sont obtenues.

4.3.2.1.1. La transition formantique et le « locus virtuel »

La production des sons de la parole nécessite le déplacement continu des organes de la parole ; déplacement qui implique d'importantes modifications de la forme et du volume du conduit vocal. Ainsi, ces variations articulatoires conduisent acoustiquement à des variations rapides visibles sur les formants lorsqu'on produit une suite syllabique de type CV ou CVC, par exemple. Le passage d'une consonne à une voyelle par exemple va modifier la forme des formants vocaliques en produisant des petits changements appelés « transitions formantiques », qui peuvent être considérées comme des variations dynamiques rapides dans la structure des formants F_1, F_2, F_3, F_4 , en fonction du temps.

Pierre Delattre et ses collaborateurs (voir Couper *et al.*, 1952 ; Delattre *et al.*, 1952, 1955) ont proposé des expériences pour étudier les variations des transitions formantiques. À partir de stimuli synthétiques, ces chercheurs ont trouvé que les auditeurs étaient capables d'identifier le lieu d'articulation des consonnes occlusives orales voisées et non voisées et les

consonnes nasales, lorsqu'un pattern de deux formants était associé à une transition formantique spécifique prenant origine à une valeur précise. Ils ont montré que lorsque les transitions formantiques de F_2 dans une syllabe de type CV étaient montantes et ayant pour origine 720 Hz, la « meilleure » consonne /b/ était identifiée par les auditeurs ; lorsque la transition partait de 1800 Hz, la « meilleure » consonne /d/ est obtenue ; et pour la « meilleure » consonne /g/, la transition était descendante de la consonne à la voyelle pour atteindre les 3000 Hz (Delattre *et al.*, 1955 : p. 771), (voir Figure 39). Ils ont utilisé le terme « locus » pour rendre compte de cette valeur fixe, point virtuel origine des transitions.

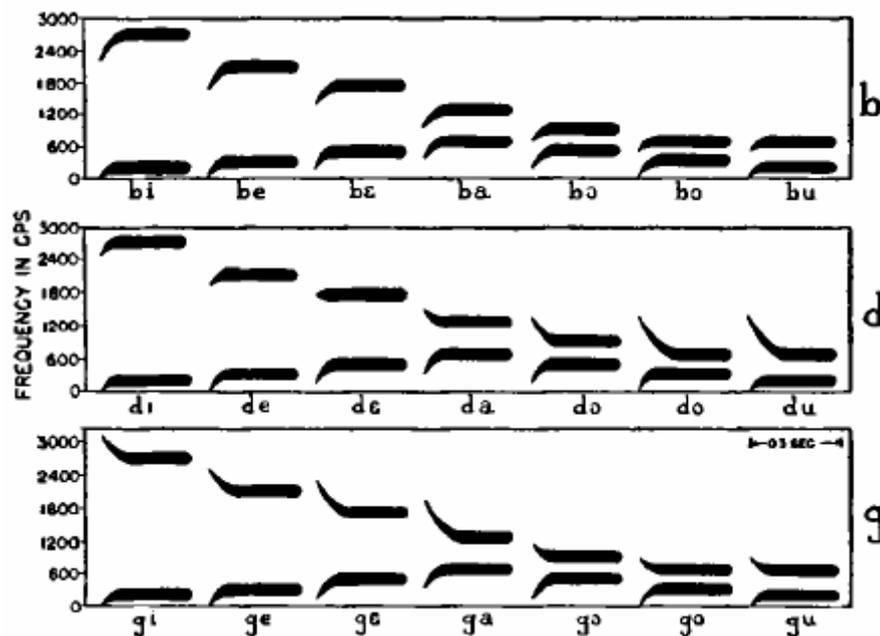


Figure 39 : Direction des transitions de F_1 & de F_2 synthétisées pour différentes voyelles montrant les lieux d'articulation correspondant. (Emprunté à Delattre *et al.*, 1955).

Afin d'évaluer l'origine de ce « locus » en fonction de la transition formantique, Delattre *et al.* (1955) ont synthétisé différents patterns de deux formants avec un état stable différent et une transition partant de 1800 Hz (pour le lieu de la consonne /d/) avec deux conditions différentes : a) la transition formantique commençait toujours à 1800 Hz et b) la transition commençait à 50 ms après le « locus », (voir Figure 40). Les auteurs constatent que lorsque la transition formantique commençait toujours à 1800 Hz (Figure 40, partie A), les auditeurs identifiaient trois lieux d'articulation au lieu d'un seul : dans la zone vocalique située entre 2040 et 2520 Hz, la consonne /b/ est identifiée ; entre 1560 à 1920, c'est la

consonne /d/ qui est identifiée ; entre 1200 et 1440 Hz, c'est la consonne /g/ qui est identifiée et enfin entre 720 et 1200 Hz, c'est à nouveau la consonne /d/ qui est identifiée. Ils constatent ensuite que lorsque les premières 50 ms après le « locus » sont effacées (Figure 40, partie B), tous les auditeurs identifiaient la consonne /d/ pour tous les stimuli allant de 720 à 2520 Hz⁴⁹. Étant donné que ce « locus » est positionné à 50 ms avant la transition formantique, on peut désormais utiliser le terme « locus virtuel » pour en rendre compte.

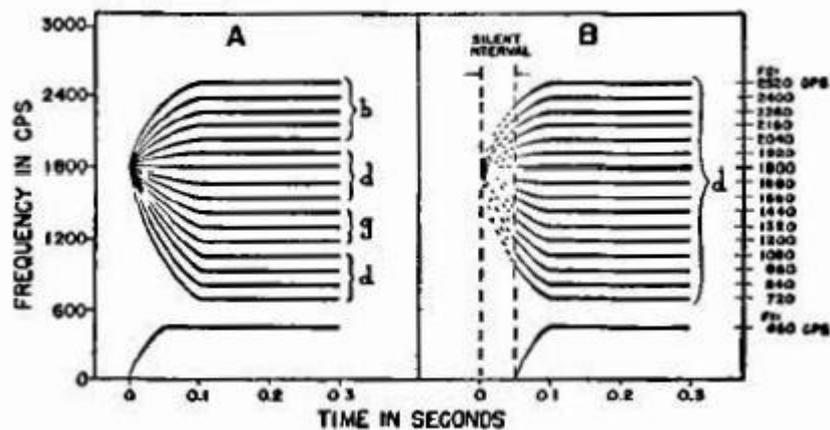


Figure 40 : Stimuli synthétisés avec les deux conditions : transition commençant au « locus » (partie A), et transition commençant à 50 ms après le « locus » (partie B), avec les catégories identifiées par les auditeurs. (Emprunté à Delattre *et al.*, 1955).

On sait depuis, que ce sont les transitions formantiques de F_2 et de F_3 , qui jouent un rôle déterminant pour indiquer le lieu d'articulation des consonnes adjacentes aux voyelles : ainsi, c'est la valeur du « locus virtuel » de F_2 et la direction descendante ou montante de F_3 qui donne une indication du lieu d'articulation. On sait également pour le lieu vélaire que ce sont deux valeurs de « locus virtuel » distinctes qui sont proposées en fonction de la voyelle adjacente : si la voyelle est antérieure, alors un « locus virtuel » autour des 3000 Hz (au maximum) est obtenu ; en revanche, si la voyelle est postérieure, alors un « locus virtuel » autour des 1300 Hz est obtenu (voir Figure 41). (Pour plus détails, voir les travaux de Kewley-Port, 1982 ; de Lindblom, 1990 ; entre autres).

⁴⁹ Les auteurs expliquent que seule une durée de 50 ms produisaient la consonne /d/ dans toutes les positions.

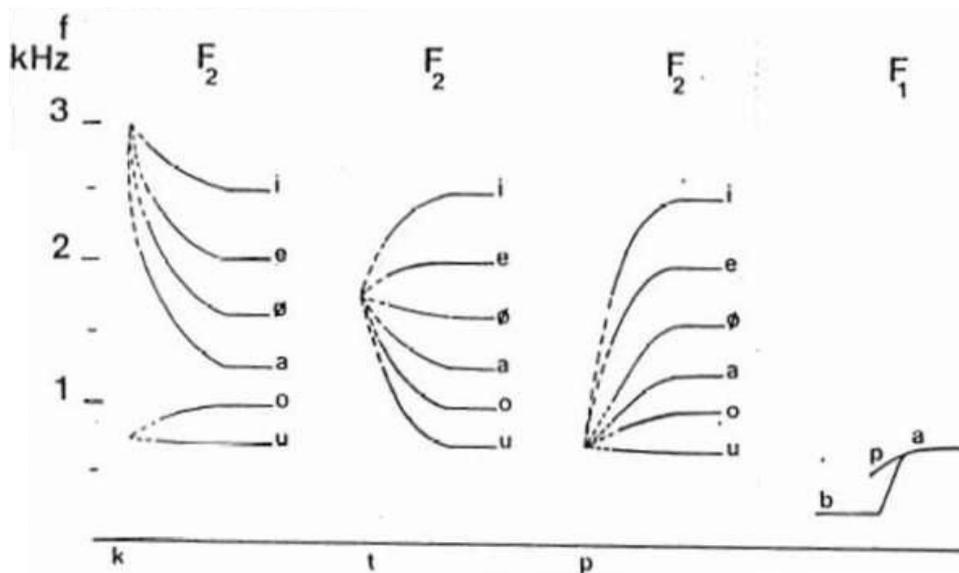


Figure 41 : Directions des transitions formantiques dans des syllabe CV avec la position du « locus virtuel » de F_2 en fonction des consonnes /p t k/ et des voyelles /i e ø a o u/. (Emprunté à Landercy & Renard, 1977).

Delattre *et al.* (1955) ont proposé une interprétation articulatoire de ce « locus virtuel » : ils le considèrent comme correspondant à des configurations invariantes du conduit vocal, point de départ de l'occlusion. Cette interprétation a été rejetée par certains chercheurs, car ces résultats ont été obtenus à partir de stimuli synthétiques et surtout en perception de la parole, mais lorsque des analyses acoustiques de la parole spontanée sont proposées, ce concept n'est pas validé (voir par exemple Kewley-Port, 1982, entre autres, qui explique que les valeurs de « locus virtuel » varient tellement pour F_2 et F_3 en fonction des voyelles environnantes, qu'on ne peut pas proposer un seul « locus virtuel » par lieu d'articulation).

Il est à noter néanmoins que l'utilisation du « locus virtuel » reste le seul moyen fiable, en plus du « burst », pour évaluer le lieu d'articulation des consonnes adjacentes (pour plus de détails sur le rôle du « burst » dans l'identification du lieu d'articulation, voir Stevens & Blumstein, 1978 ; Blumstein & Stevens, 1979, 1980 ; entre autres).

À partir de cette notion de « locus virtuel », différents auteurs ont tenté d'évaluer le rôle des transitions formantiques dans la détermination du lieu d'articulation des consonnes adjacentes à partir des **équations de locus**.

4.3.2.1.2. Les équations de locus

Les **équation de locus** consistent en des régressions linéaires effectuées à partir de points représentatifs de la relation entre les **onsets** du deuxième formant de différentes voyelles (F_{2Onset}) et leurs états stables correspondant ($F_{2Milieu}$). Cette équation a la forme suivante : $F_{2Onset} = m * F_{2Milieu} + b$ (où m et b sont respectivement la pente et l'ordonnée à l'origine de l'équation de régression). Les valeurs de cette équation sont obtenues pour une seule consonne coarticulée avec différentes voyelles, où F_{2Onset} représente l'ordonnée et $F_{2Milieu}$ l'abscisse (pour un exemple tiré de nos données, voir Figure 42). Björn Lindblom est à l'origine du concept de l'**équation de locus** (Lindblom, 1963_b), à partir d'une étude sur des séquences CV (où C = /b d g/ suivie des 8 voyelles du suédois) sur un seul sujet. Les coefficients des **équations de locus** obtenus varient en fonction du lieu d'articulation. Ainsi, il a trouvé que les pentes de la droite de régression variaient en fonction du lieu d'articulation de la consonne, de /g/ (0,95) > /b/ (0,69) > /d/ (0,28). L'auteur ne donne aucune interprétation de ces résultats mais explique que "*Fitting straight lines serves only a descriptive purpose and is not motivated by deeper considerations as to the underlying complex articulatory processes*" (Lindblom, 1963_b : p. 67). L'auteur utilise la notion « locus » pour désigner une valeur mesurée concrète, à la place d'une valeur dérivée par le « locus virtuel » ; la valeur de F_{2Onset} qui varie systématiquement avec la voyelle adjacente. À partir de ces résultats, on peut facilement identifier le lieu d'articulation des consonnes, i.e., le « locus virtuel », en fonction à la fois de la valeur de la pente de la droite de régression et de l'ordonnée à l'origine. La valeur du « locus virtuel » correspond au point d'intersection de la ligne de régression avec la diagonale à 45 degrés (ou en appliquant la formule suivante : « locus virtuel » = $b / (1 - m)$).

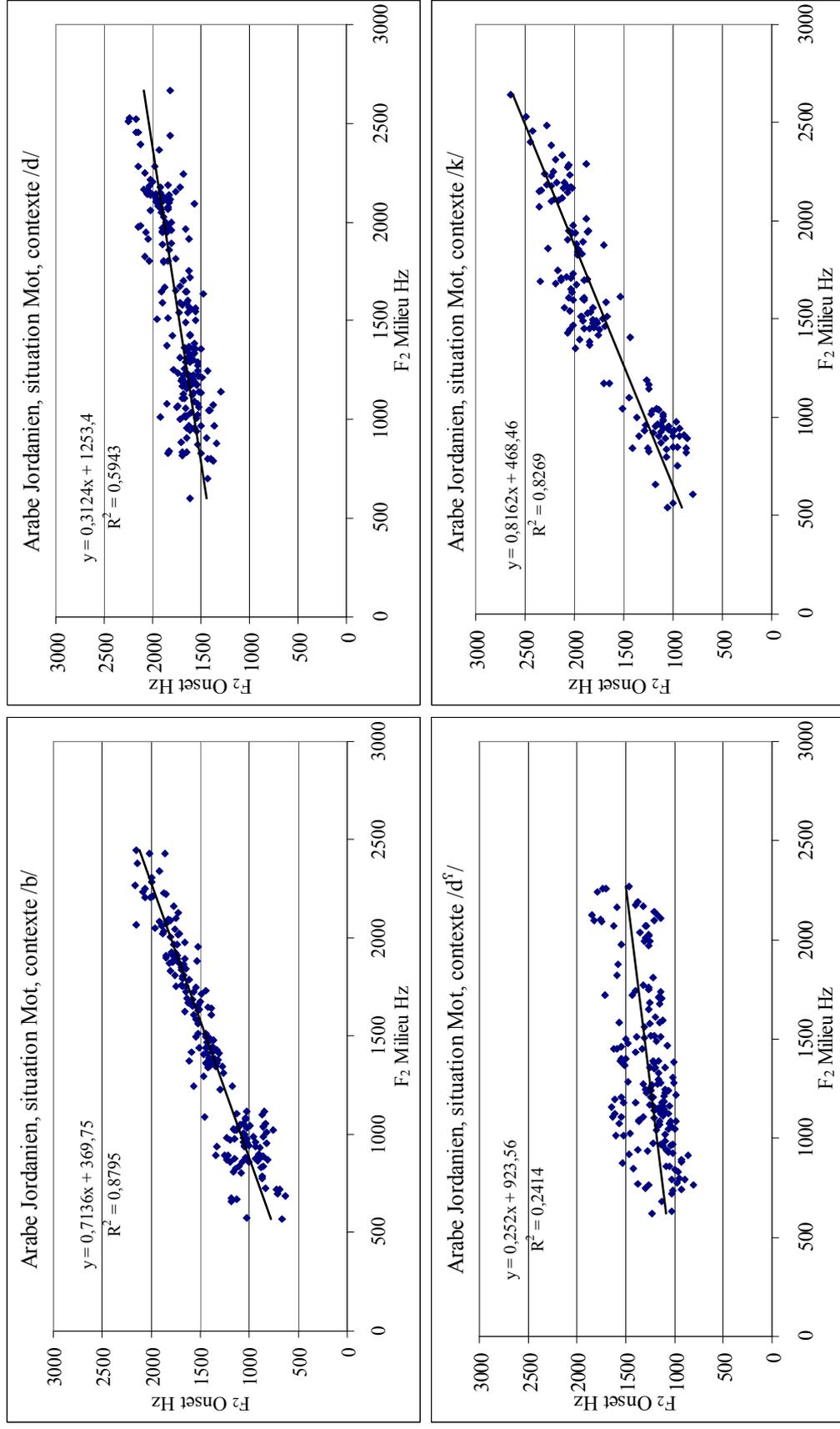


Figure 42 : Représentation en nuage de points de la relation entre F_{2Onset} et $F_{2Milieu}$ et la droite de régression pour les consonnes /b d^v k/ de la coarticulation CV, avec les coefficients de l'équation de locus. Les points indiquent des voyelles.

Depuis l'étude de Lindblom (1963_b), plusieurs chercheurs ont exploité cette méthode afin d'identifier le lieu d'articulation des consonnes, nous pouvons citer entre autres : Krull, 1987, 1988 ; Sussman *et al.*, 1991, 1993, 1997 ; Fowler, 1994 ; Celdran & Villabla, 1995 ; Sussman & Shore, 1996 ; Chennoukh *et al.*, 1997 ; Yéou, 1995, 1997 ; Tabain & Butcher, 1999 ; Tabain, 2000 ; Sussman & Modarresi, 2003 ; Al-Tamimi, 2004 ; Modarresi *et al.*, 2005 ; etc...). Toutefois, certains auteurs ont trouvé des différences pour les coefficients des **équations de locus** pour des consonnes ayant le même lieu d'articulation et différant par le mode d'articulation et/ou le voisement (voir Krull, 1987, 1988 ; Fowler, 1994 ; Chennoukh *et al.*, 1997 ; Yéou, 1995, 1997 ; etc...). Ils considèrent ainsi les coefficients obtenus comme de bons indicateurs du degré de coarticulation des voyelles avec les consonnes adjacentes : une valeur de pente de la droite de régression élevée (i.e., $m = 1$) indique une coarticulation maximale entre consonnes et voyelles (i.e., une résistance minimale du geste consonantique aux effets coarticulatoires des voyelles), tandis qu'une pente faible (i.e., $m = 0$) indique l'absence de coarticulation entre consonnes et voyelles (i.e., une résistance maximale du geste consonantique aux effets coarticulatoires des voyelles).

Nous avons voulu savoir si les valeurs de la pente de la droite de régression obtenues dans les différentes études citées plus haut sont homogènes entre elles ou non. Nous présentons ainsi dans le Tableau 10 les valeurs de la pente de la droite de régression des **équations de locus** pour F_2 , obtenues pour les trois lieux d'articulation bilabial, dental et vélaire pour plusieurs langues ayant différents systèmes vocaliques.

Tableau 10 : Valeurs moyennes de pentes de la droite de régression des **équations de locus**, en fonction des consonnes /b d g/⁵⁰, en contexte CV, à partir de différentes études. (* = donnée non disponible ou contexte non traité). Si le contexte consonantique est différent, nous l'avons marqué entre parenthèse.

	Langue	/b/	/d/	/g/ (vélaire)
Lindblom (1963 _b)	Suédois	0,69	0,28	0,95
Krull (1987)	Suédois	0,78	0,40	*
Sussman <i>et al.</i> (1991)	Anglais Américain	0,89	0,42	0,70
Sussman <i>et al.</i> (1993)	Thai	0,70	0,30	*
Sussman <i>et al.</i> (1993)	Arabe égyptien	0,77	0,25	0,92
Sussman <i>et al.</i> (1993)	Urdu	0,81	0,50	0,97
Fowler (1994)	Anglais Américain	0,80	0,48	0,71
Celdran & Villabla (1995)	Espagnol	0,83	0,58	1,00
Sussman & Shore (1996)	Anglais Américain	*	0,40	*
Chennoukh <i>et al.</i> (1997)	Ph1 : coarticulation maximale	0,85	0,51	0,98
Chennoukh <i>et al.</i> (1997)	Ph2 : coarticulation moyenne	0,75	0,42	0,86
Chennoukh <i>et al.</i> (1997)	Ph3 : coarticulation minimale	0,33	0,22	0,45
Sussman <i>et al.</i> (1997)	Anglais Américain	0,77	0,37	1,12
Yéou (1997)	Arabe Standard	0,92 (/f/)	0,48	0,90 (/χ/)
Tabain & Butcher (1999)	Yanyuwa	0,80 (/p/)	0,62 (/t/)	1,03 (/k/)
Tabain & Butcher (1999)	Yindjibarndi	0,83 (/p/)	0,48 (/t/)	1,11 (/k/)
Tabain (2000)	Anglais Australien	*	0,42	0,84
Sussman & Modarresi (2003)	Anglais Américain	0,80	0,41	0,89
Sussman & Modarresi (2003)	Perse	0,87	0,48	1,03
Al-Tamimi (2004)	Arabe Jordanien	0,63	0,36	0,87
Modarresi <i>et al.</i> (2005)	Anglais Américain	0,69 (@burst)	0,29 (@burst)	0,94 (@burst)
Modarresi <i>et al.</i> (2005)	Perse	0,80 (@burst)	0,30 (@burst)	1,00 (@burst)

Nous observons des différences de valeurs pour ces trois lieux, qui sont très probablement dues aux différences liées au nombre de consonnes et voyelles dans chaque langue, mais également aux méthodes d'extraction de l'**onset** vocalique (à la première pulsation ou période vocalique, @burst⁵¹, etc... (voir Figure 43)). Toutefois, nous observons une tendance : en moyenne, les valeurs des pentes des **équations de locus** sont élevées pour les consonnes vélaïres, intermédiaires pour les bilabiales et basses pour les dentales (avec des valeurs de pentes de la droite de régression moyennes de 0,93, 0,79 & 0,42, respectivement). Ceci indique un degré de coarticulation important pour les vélaïres, intermédiaires pour les

⁵⁰ Il est à noter que pour le lieu vélaire, tous les auteurs ont expliqué la nécessité d'utiliser deux **équations de locus** différentes pour rendre compte des différences liées aux voyelles adjacentes. Ainsi, nous obtenons deux valeurs de pente de la droite de régression et deux ordonnées à l'origine.

⁵¹ @burst est une mesure effectuée directement après la barre d'explosion (Modarresi *et al.*, 2005). Deux points de mesure de l'**onset** vocalique sont proposés dans la littérature : soit à partir de la première pulsation vocalique (i.e., sans prendre en compte la barre d'explosion ni le bruit de friction engendré) soit directement après la barre d'explosion (et par conséquent intégrant le bruit de friction dans la transition vocalique).

bilabiales et moins important pour les dentales. Yéou (1995, 1997) explique que la différence des valeurs de pente obtenues entre les lieux d'articulation peut être corrélée à l'utilisation de la langue comme articulateur principal dans la production des consonnes. En effet, les différentes études portant sur la coarticulation montrent que le contrôle de cet articulateur est inversement proportionnel au degré de coarticulation. Puisque la langue n'intervient pas dans d'articulation d'une consonne bilabiale, les effets de coarticulation consonne-voyelle sont intermédiaires. En revanche, la production d'une dentale nécessite la langue comme articulateur principal, permettant ainsi une résistance « maximale » de la coarticulation consonne-voyelle. Quant à l'articulation d'une consonne vélaire, l'utilisation du dos de la langue ne permet pas toute la précision articulatoire nécessaire pour permettre au geste consonantique de résister aux effets coarticulatoires des voyelles et vice-versa et par conséquent la valeur de sa pente est très élevée.

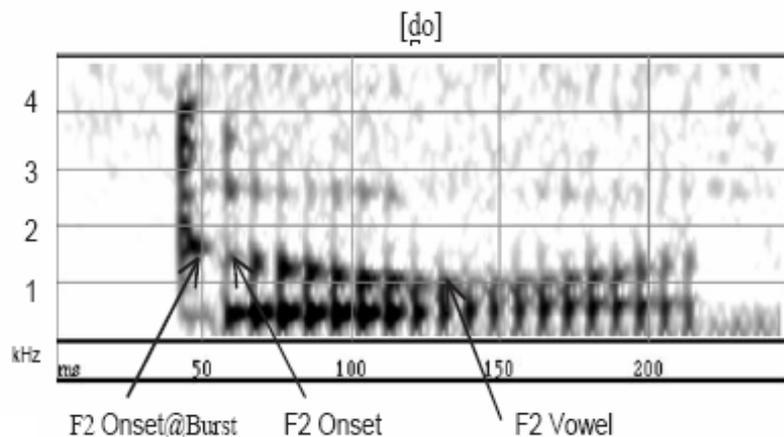


Figure 43 : Schéma présentant les méthodes d'extraction de l'**onset** : a) @Burst et b) à la première pulsation après le relâchement de la voyelle. (Emprunté à Modarresi *et al.*, 2005).

Cette présentation « quasi dynamique » des transitions formantiques rend compte uniquement des liens existant entre le lieu d'articulation des consonnes adjacentes et les influences observées sur les valeurs formantiques de l'état stable et de l'**onset** vocalique. Si nous voulons quantifier la valeur du « locus virtuel » ou celle de la pente de la droite de régression pour une seule voyelle, le peu de données ne permet pas d'obtenir de mesures fiables. Il faut ajouter à cela que, dans notre travail, nous cherchons à identifier les paramètres

dynamiques qui permettent différencier les voyelles entre elles. C'est pourquoi nous avons adapté le calcul de l'**équation de locus** aux voyelles.

4.3.2.1.3. Les pentes formantiques

Le calcul des **pentés formantiques** que nous proposons est une version améliorée de l'**équation de locus**. La différence entre les deux méthodes est que les **équations de locus**, qui sont obtenues à partir de deux valeurs formantiques obtenues à F_{2Onset} et à $F_{2Milieu}$, donnent une bonne description du lieu d'articulation et/ou du degré de coarticulation ; tandis que les **pentés formantiques** qui sont obtenues à partir de plusieurs valeurs formantiques (au minimum cinq valeurs pour une durée de transition de 20 ms), rendent compte du degré de coarticulation et du lieu d'articulation des consonnes adjacentes pour chaque voyelle et chaque formant. La **pente formantique** calculée ici est plus proche de la vitesse de transition formantique. Comme nous l'avons expliqué, l'**onset** vocalique est mesuré à 5 ms après le relâchement de la voyelle ; position située toujours au milieu de la première période de la voyelle. Dès que l'**onset** est déterminé, toutes les valeurs obtenues à 5 ms d'intervalle sont utilisées pour effectuer le calcul des **pentés formantiques**. Ainsi, et pour tous les formants ; les pentes formantiques sont obtenues par une analyse de régression linéaire en partant de la première valeur formantique, obtenue à 5 ms du début, qui est placée au temps 0, jusqu'à la valeur au milieu temporel de la voyelle. Nous avons décidé de nommer l'intervalle de temps entre le point 0 et le milieu temporel la « durée de transition », qu'il ne faut cependant pas confondre avec la terminologie traditionnelle. La durée de la transition est traditionnellement quantifiée à partir du point 0 (début de la voyelle) jusqu'à la fin des changements abruptes dûs au passage de la consonne à la voyelle. Kent & Moll (1969) expliquent que la durée de la transition vers la voyelle en contexte CV est un invariant **dynamique** en production des occlusives voisées et non-voisées et correspond à une durée proche des 50 ms. Au-delà de ces 50 ms, l'état stable de la voyelle débute. Dans notre cas, nous prenons en compte la durée de la transition, traditionnelle, mais également une partie de la durée de l'état stable. Pour toutes les voyelles ayant une durée vocalique inférieure à 100 ms, la durée de la transition traditionnelle et celle que nous utilisons seront les mêmes, à partir des suppositions de Kent &

Moll (1969) ; pour les voyelles ayant une durée vocalique supérieure à 100 ms, la transition traditionnelle est prise en compte avec en plus une partie de l'état stable.

Dès lors que tous les paramètres nécessaires à effectuer le calcul des **pent**es **formantiques** sont réunis, nous utilisons une analyse de régression linéaire (effectuée à partir d'un script que nous avons élaboré sous le logiciel Praat) – en suivant la même formule utilisée précédemment : $y = m*x + b$, où les valeurs de x représentent le temps et les valeurs y les formants – en partant de la valeur à l'**onset** jusqu'à celle du milieu temporel de la voyelle (voir Figure 44 pour plus de détails concernant le calcul de régression linéaire).

Ce calcul donne une bonne caractérisation **dynamique** des influences consonantiques sur les voyelles, car directement affecté par les trajectoires réelles des transitions formantiques. Si la durée de transition est inférieure à 50% de la durée vocalique ; nous obtenons des mesures fidèles, en revanche, si la durée de la transition est supérieure à 50% de la durée vocalique, des mesures moins représentatives sont obtenues du fait de l'intégration d'une partie importante de la durée de l'état stable dans le calcul.

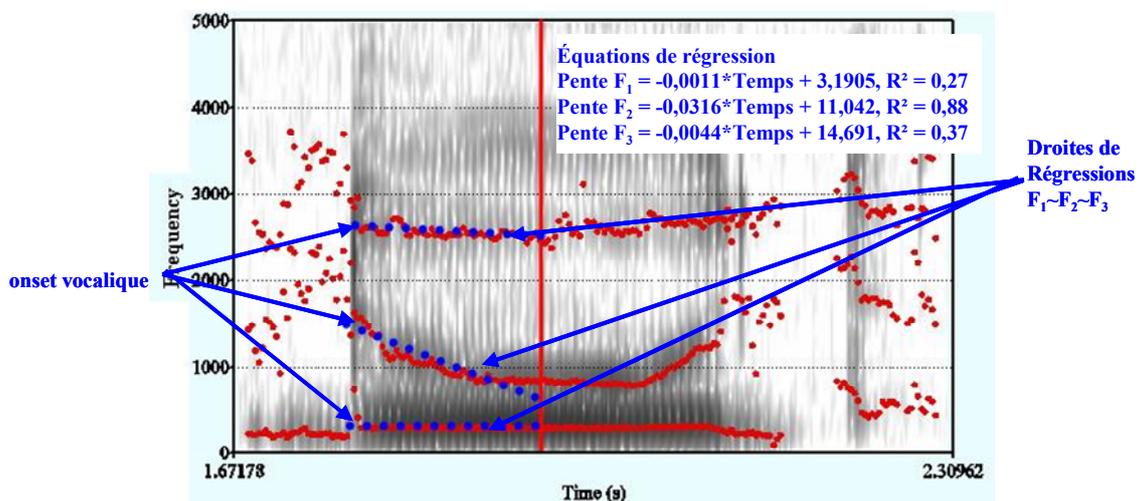


Figure 44 : Schéma représentant le calcul de régression linéaire pour la voyelle /u:/ comme réalisée dans le mot /du:d/ « vers de terre (pl.) » par le locuteur 09MAJ, avec la position de l'onset vocalique, ainsi que les droites et les équations de régression linéaire.

Le calcul de la **pen**te **formantique** que nous proposons est linéaire ; dans les situations de production Mot et Syllabe, toutes les trajectoires formantiques ne sont pas prises en compte, car les transitions sont souvent courbées et ne suivent pas une ligne droite.

L'exemple présenté sur la Figure 44 rend compte de ce problème et plus particulièrement pour la **pen**te formantique de F_2 qui n'est pas bien modélisée : la transition est non linéaire et les valeurs formantiques ne suivent pas la droite de régression. C'est pourquoi, nous avons effectué deux autres calculs de régression non-linéaire : régressions polynomiales de 2^{ème} (quadratique) et de 3^{ème} (Cubique) ordre⁵² (effectuées avec le logiciel MATLAB), où les valeurs de x représentent le temps, et les valeurs de y , les formants, qui sont représentées par les formules suivantes :

$$y = a_0 + a_1x + a_2x^2, \text{ régression polynomiale de } 2^{\text{ème}} \text{ ordre (Quadratique),}$$

$$y = a_0 + a_1x + a_2x^2 + a_3x^3, \text{ régression polynomiale de } 3^{\text{ème}} \text{ ordre (Cubique).}$$

Nous présentons sur la Figure 45 les droites de régressions des trois analyses **dynamiques** appliquées sur nos données : la régression linéaire et les deux régressions polynomiales de 2^{ème} et de 3^{ème} ordre, pour la voyelle /u:/ comme réalisée dans le mot /du:d/ « vers de terre (pl.) » par le locuteur 09MAJ. Dans cet exemple, nous observons que les trajectoires de F_1 et de F_3 sont bien modélisées à partir de la régression linéaire ; l'augmentation de l'ordre (i.e., en utilisant la régression polynomiale de 2^{ème} ou de 3^{ème} ordre) n'améliore pas la modélisation de ces trajectoires, car elles suivent une ligne droite. La modélisation de la trajectoire de F_2 en revanche est améliorée en utilisant la régression polynomiale de 3^{ème} ordre : en augmentant l'ordre, la valeur de coefficient de corrélation (R^2) augmente et se rapproche de 1. Nous observons qu'en fonction de chaque analyse de régression, la valeur de départ de transition formantique (ou l'**onset** dérivé) est différente et que par exemple pour la transition de F_2 , c'est la régression polynomiale de 3^{ème} qui arrive à dériver un **onset** proche de celui réellement extrait. À partir de cet exemple, nous pouvons dire que l'utilisation de la régression linéaire est suffisante pour F_1 et F_3 , tandis que c'est la régression polynomiale de 3^{ème} ordre qui améliore la modélisation de la trajectoire de F_2 .

⁵² Les résultats des deux analyses de régressions polynomiales seront utilisés uniquement lors de l'analyse discriminante car il est difficile de faire une représentation multidimensionnelle des résultats obtenus. Les corrélats linguistiques directs de ces coefficients sont les suivants : pour la régression quadratique, a_0 = la moyenne des valeurs, a_1 = la droite de régression et a_2 = l'ordonnée à l'origine ; pour la régression cubique, a_0 = la moyenne des valeurs, a_1 = la droite de régression, a_2 = l'excursion parabolique de la voyelle indépendamment de sa droite et a_3 = l'ordonnée à l'origine, (McDougall, 2006).

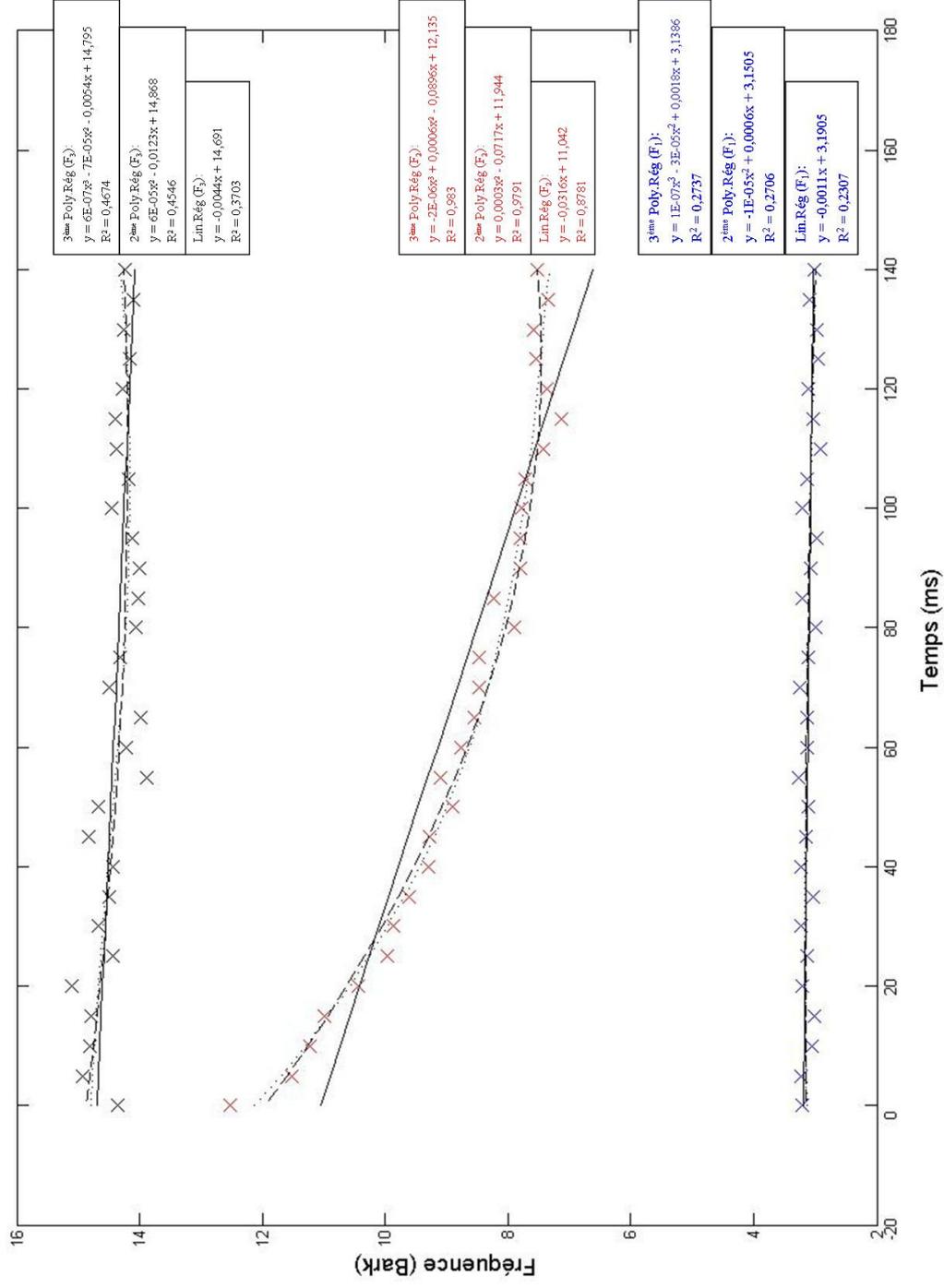


Figure 45 : Droites de régression linéaires (en lignes continues), polynômes de 2^{ème} ordre (en lignes discontinues) et polynômes de 3^{ème} ordre (en lignes pointillées) de F₁, F₂ & F₃ de la voyelle /u:/ comme réalisée dans le mot /dur/ « vers de terre (pl.) » par le locuteur 09MAJ, ainsi que les équations de régression.

Avec ces deux analyses polynomiales, nous nous attendons à observer une amélioration de la classification des voyelles dans certains contextes consonantiques. Nous évaluerons ainsi laquelle des trois analyses est pertinente pour une « meilleure » description des voyelles dans les trois langues étudiées.

Afin d'interpréter nos résultats, nous nous servons des prédictions formulées dans la littérature concernant le « locus virtuel » et les **équations de locus**. Ainsi, le lien entre le degré de coarticulation (i.e., valeur de pente de la droite de régression) et le degré de la réduction vocalique dû aux effets du lieu d'articulation des consonnes adjacentes peut être expliqué suivant la relation linéaire entre F_{2Onset} et $F_{2Milieu}$: les modifications de valeurs de $F_{2Milieu}$ affecteront celles de F_{2Onset} et par conséquent celles des valeurs de pentes des **équations de locus**. Prenons l'exemple de la voyelle /i/ : lorsque cette voyelle est réduite vers [ə], la valeur de son $F_{2Milieu}$ va être plus basse et va directement affecter la valeur de F_{2Onset} en l'abaissant, ainsi la valeur de pente de la droite de régression de l'**équation de locus** devient plus basse (i.e., se rapprochant de 0). Nous donnons dans Figure 46 une explication schématique de cet exemple. Ainsi, dans cet exemple, trois syllabes /di/ sont schématisées : la première présente une voyelle /i/ pleine non réduite en ligne noire (avec des fréquences de $F_1 = 250$ Hz ; $F_2 = 2300$ Hz ; $F_3 = 3000$ Hz) ; la deuxième présente une voyelle /i/ pleine légèrement réduite en ligne rouge (avec des fréquences de $F_1 = 300$ Hz ; $F_2 = 2200$ Hz ; $F_3 = 2800$ Hz) ; la troisième présente une voyelle /i/ très réduite en [ə] en ligne bleue (avec des fréquences de $F_1 = 500$ Hz ; $F_2 = 1500$ Hz ; $F_3 = 2500$ Hz). Dans les trois syllabes, les transitions formantiques pointent toutes vers le lieu d'articulation /d/ avec un « locus virtuel » à 1800 Hz (comme proposé par Delattre *et al.*, 1955). Les valeurs des « cibles vocaliques » de ces trois voyelles sont différentes en fonction du degré de réduction vocalique, et ces variations de fréquences à l'état stable affectent les valeurs des **onsets** vocaliques obtenues au moment du relâchement de la voyelle (présenté dans le graphique en ligne verticale bleue en pointillé). Ce sont donc les modifications observées pour les fréquences de l'état stable et de l'**onset** vocalique qui vont modifier les valeurs des **pentés formantiques** dérivées pour chaque formant. Avec les « locus virtuel » et les **équations de locus**, on ne tiendra pas compte de ces modifications, car elles sont gommées par la généralisation effectuée.

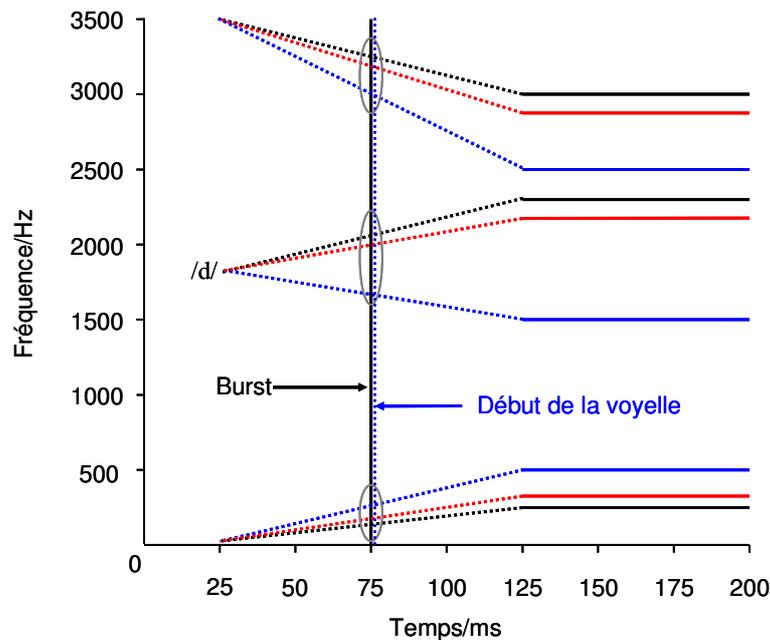


Figure 46 : Syllabe /di/ schématisée en trois conditions : voyelle /i/ pleine non réduite (en ligne noire) ; voyelle /i/ pleine légèrement réduite (en ligne rouge) et voyelle /i/ réduite en /ə/ (en ligne bleue). Les cercles entourent les **onsets** vocaliques extraits juste après le début de la voyelle.

Par conséquent, nous pensons que les valeurs de **pentés formantiques** obtenues par régression linéaire sont directement corrélées au degré d'inclinaison des transitions formantiques : à durée égale, **une inclinaison forte des transitions formantiques pour n'importe quel formant donnerait des valeurs de pentés élevées**. Nous nous attendons donc à observer des différences liées au lieu d'articulation, à la voyelle et au formant choisi. En suivant les prédictions des **équations de locus** et les résultats du « locus virtuel » (Delattre *et al.*, 1955), nous pouvons supposer que :

- Les valeurs des **pentés formantiques** dépendent du lieu d'articulation : nous obtiendrons des valeurs de **pentés** élevées en contexte dental, par rapport au contexte bilabial et basses en contexte vélaire⁵³,
- Les valeurs des **pentés formantiques** dépendent des voyelles : sur F_1 , /a/ obtient les valeurs de **pentés** les plus élevées, /i/ et /u/, les plus basses, les autres voyelles sont entre les deux ; sur F_2 , les voyelles postérieures, et plus spécifiquement /u/,

⁵³ Ces suppositions peuvent être explicitées par le fait que les mouvements de la pointe de la langue sont plus rapides que ceux des lèvres et enfin les mouvements obtenus pour le dos de la langue sont les plus lents.

obtiennent les valeurs de **pentés** les plus importantes, les moins importantes pour les voyelles antérieures, et plus spécifiquement /i/ ; /a/ se situe entre les deux.

Afin de valider nos résultats, nous avons effectué une analyse statistique : une MANOVA à quatre facteurs (langue (3), situations de production (3), lieux d'articulation (quatre en AM et AJ et trois en FR) et voyelles (cinq en AM, huit en AJ et onze en FR)).

Dans la partie qui suit, nous présentons les résultats de la comparaison inter-langues en **dynamique** pour les trois voyelles communes /i a u/ (voyelles longues en arabe), afin de valider les effets de la densité des systèmes observés en **statique**.

4.3.2.2. Résultats des pentés formantiques

Nous comparons dans cette partie les trois voyelles communes /i a u/ dans les trois langues en **dynamique**. Rappelons que les résultats de la comparaison inter-langues en **statique** montrent globalement une tendance à la centralisation de l'espace vocalique en AM par rapport à celui en AJ ou en FR. Nous examinerons donc ces effets en fonction du degré d'inclinaison des pentés formantiques.

La durée de la transition formantique est significativement différente en fonction de la langue : en moyenne, elle est plus courte en FR et plus longue en AM ($F(3, 4768)=296,17$; $p<0,001$). L'interaction entre la langue et la situation de production s'est révélée significative : en AM et en situation Mot, la durée de la transition formantique est la plus longue par rapport au FR en situation Mot qui présente la durée la plus courte ($F(6, 4768)=57,31$; $p<0,001$). Toutes les autres interactions entre les langues, la situation de production, le lieu d'articulation et les voyelles se sont révélées significatives ($p<0,001$).

Nous observons que la différence moyenne entre la durée de la transition formantique obtenue dans les deux dialectes arabes et le FR est de l'ordre de 50 ms. La littérature concernant l'étude de la réduction vocalique explique que l'utilisation du paramètre de la durée semble affecter le degré de réduction vocalique dans les langues (voir Lindblom, 1963_a). Toutefois, certains auteurs ont montré que la durée vocalique n'a pas de rôle important dans le phénomène de la réduction vocalique (voir Gay, 1978 ; Pols & van Son,

1993). Les données de la littérature sont donc contradictoires. Nous nous attendons à observer deux cas de figures : soit la durée influence le degré de réduction vocalique, soit elle n'intervient pas. Si la durée vocalique affecte le degré de réduction vocalique, nous nous attendons à observer son influence directement sur nos données : plus la durée vocalique est longue, plus la cible vocalique est atteinte et moins l'inclinaison de la transition formantique est forte. À durée variable (comme entre les voyelles longues de l'arabe et "brèves" du FR, voir Tableau 11), nous devons nous attendre alors à observer des valeurs de **pentés formantiques** plus élevées en FR et plus basses en AM ou en AJ (i.e., degré d'inclinaison plus forte en FR et plus faible en AM ou en AJ). Afin d'évaluer le rôle de la durée de la transition formantique sur le degré d'inclinaison des **pentés formantiques**, nous avons décidé de comparer les résultats avec et sans normalisation du temps. Ainsi, en ce qui concerne les **pentés formantiques** obtenues sans normalisation du temps, les valeurs de x – dans la formule $y = m*x + b$ – correspondent aux valeurs de la durée de la transition formantique de chaque voyelle pour chaque langue. En revanche, en ce qui concerne les **pentés formantiques** obtenues avec la normalisation, les valeurs de la durée de la transition formantique ont été situées dans un intervalle de 0 à 0,5, où 0 = le début de l'onset et 0,5 = le milieu temporel de la voyelle. Autrement dit, la durée de chaque **pente formantique** sera identique lorsque le temps est normalisé. De cette façon nous apprécierons tout simplement le degré d'inclinaison de chaque **pente** indépendamment de sa longueur.

Tableau 11 : Durées moyennes des transitions formantiques (en ms) pour les voyelles /i a u/ en AM, AJ et FR.

	/i/		/a/		/u/	
	moyenne	écart type	moyenne	écart type	moyenne	écart type
Arabe Marocain	138,89	39,70	134,24	37,03	118,70	40,70
Arabe Jordanien	112,78	27,23	115,13	30,90	103,23	28,33
Français	81,57	20,51	78,18	14,09	85,04	20,40

Les résultats de la comparaison des voyelles /i a u/ dans les trois langues, en **statique** (voir § 4.3.1) montraient l'existence d'une différence importante entre les dispersions des voyelles dans les trois langues : en AM, l'espace vocalique moyen, tout contexte confondu)

est plus « centralisé » par rapport à ceux en AJ ou FR. Alors, nous nous attendons à obtenir des résultats concordant avec les deux types d'analyse (i.e., avec et sans normalisation du temps), car à notre sens, la normalisation de la durée de la transition formantique n'affectera pas le degré d'inclinaison des transitions formantiques. En effet, si la durée de la transition formantique est située dans un intervalle de 0 à 0,5, ceci aura pour effet de « réduire » simplement l'étendu de la transition et n'affectera pas son inclinaison. De cette façon, nous nous attendons à ce que le degré d'inclinaison de la transition rendra compte davantage des effets de la densité des systèmes : pour une voyelle précise, une **pente formantique** « forte » ou « raide » indique un degré de « centralisation » faible, car la production de la voyelle se rapprocherait davantage de la production d'un « cible vocalique » *Hyper-Articulée*. Nous présenterons ainsi les résultats des **pentés formantiques** sans normalisation et avec normalisation du temps et évaluerons ensuite la pertinence de ces indices à partir de la classification des voyelles des trois langues.

4.3.2.2.1. Sans normalisation du temps

Les résultats, obtenus à partir des **pentés formantiques** sans la normalisation du temps, montrent une différence importante au niveau du degré d'inclinaison des pentes formantiques moyennes entre les trois langues. Les Figures 47 à 49 montrent les **pentés formantiques** pour les trois voyelles communes /i a u/ dans les trois langues en situation de production Mot (nous avons décidé de ne présenter que les résultats obtenus dans cette situation de production car ils sont beaucoup plus marqués ; pour les résultats graphiques des **pentés formantiques** en fonction des deux autres situations, voir Annexes 2 & 3 respectivement)⁵⁴. Les **pentés formantiques** représentées sur les graphiques sont obtenues en partant de la valeur de l'ordonnée à l'origine et de la « cible vocalique » dérivée à partir de l'équation de régression. L'observation des résultats des **pentés formantiques** indique un degré d'inclinaison significativement important des transitions formantiques de F₂ seulement ; en FR, les **pentés formantiques** sont plus « raides » que celles en AM ou en AJ : sur F₂, les

⁵⁴ En situation Syllabe, nous obtenons approximativement les mêmes résultats qu'en Mot ; en Isolation en revanche, nous obtenons des **pentés formantiques** presque « plates » dans les trois langues, mais « non-nulles ». Ceci indique des trajectoires de formant propres aux voyelles.

locuteurs français produisent des voyelles beaucoup plus « périphériques » qui se rapprochent des « cibles vocaliques » et ceci est marqué par la forte inclinaison des pentes ($F(3, 4768)=3,04$; $p<0,05$). Aucune différence significative n'est obtenue sur F_1 , indiquant que l'inclinaison des **pentés formantiques** est la même dans les trois langues. Ce résultat est logique car comme nous venons de le voir, la transition de F_1 , est toujours montante de la consonne à la voyelle d'après les prédictions de la théorie du « locus virtuel ».

Les points de départ des transitions formantiques et le degré d'inclinaison des **pentés formantiques** sont différents en fonction de l'interaction entre la langue et le lieu d'articulation, uniquement sur F_2 : l'AJ et l'AM présentent les **pentés formantiques** moyennes les moins « raides » ; et le FR, les plus « raides » ($F(9, 4768)=7,36$; $p<0,001$). Comme nous l'avons expliqué auparavant, les valeurs de l'état stable influencent directement celles de l'**onset** vocalique : comme les valeurs de l'état stable sont globalement plus « centralisées » en AM sur les deux axes F_1 , F_2 , nous obtenons une « centralisation » importante du point de départ des transitions formantiques des voyelles en AM, uniquement sur F_2 ($F(3, 4768)=35,07$; $p<0,001$).

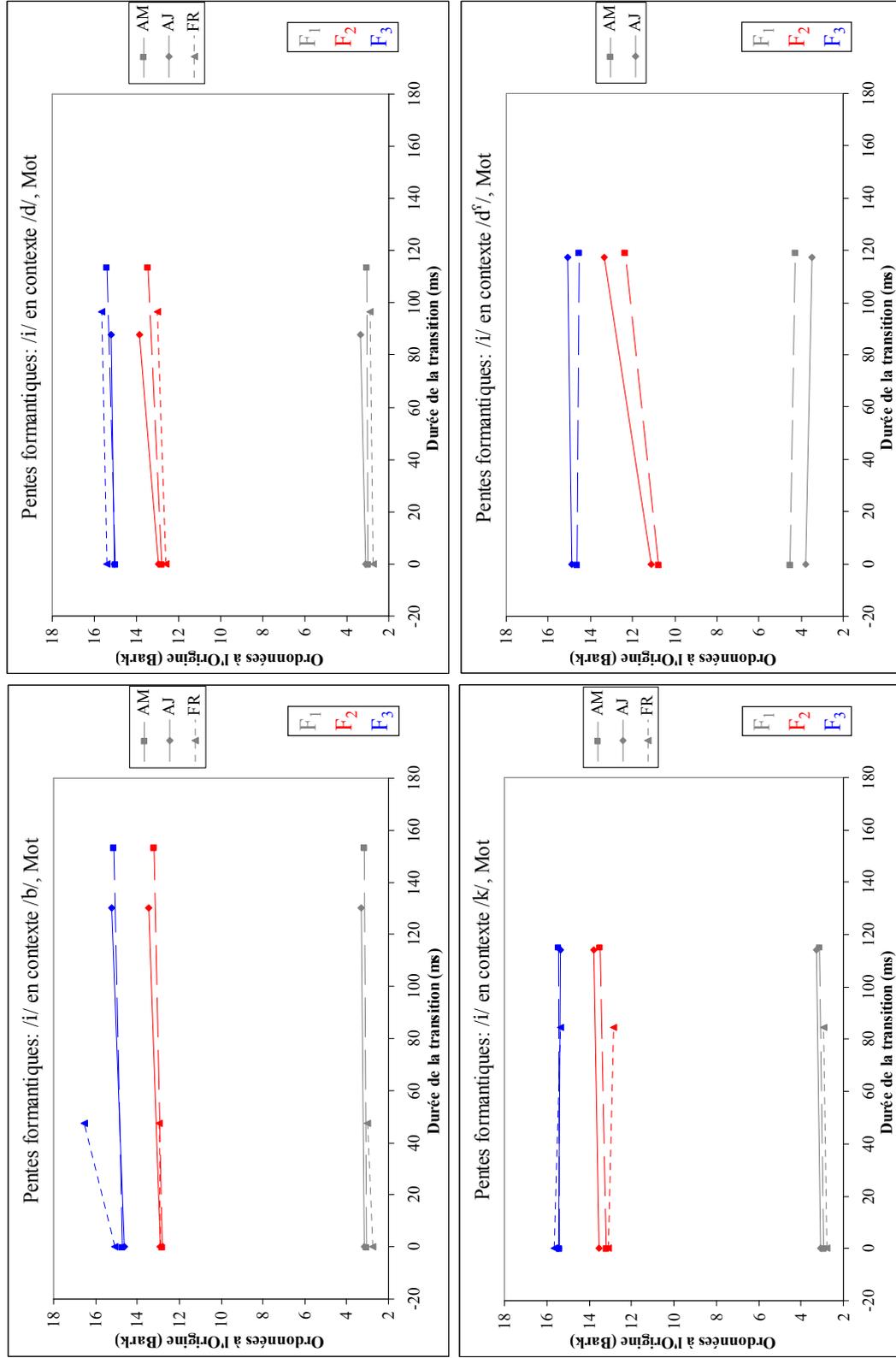


Figure 47 : Pentas formantiques de F₁, F₂ & F₃ sans normalisation du temps, pour la voyelle /i/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

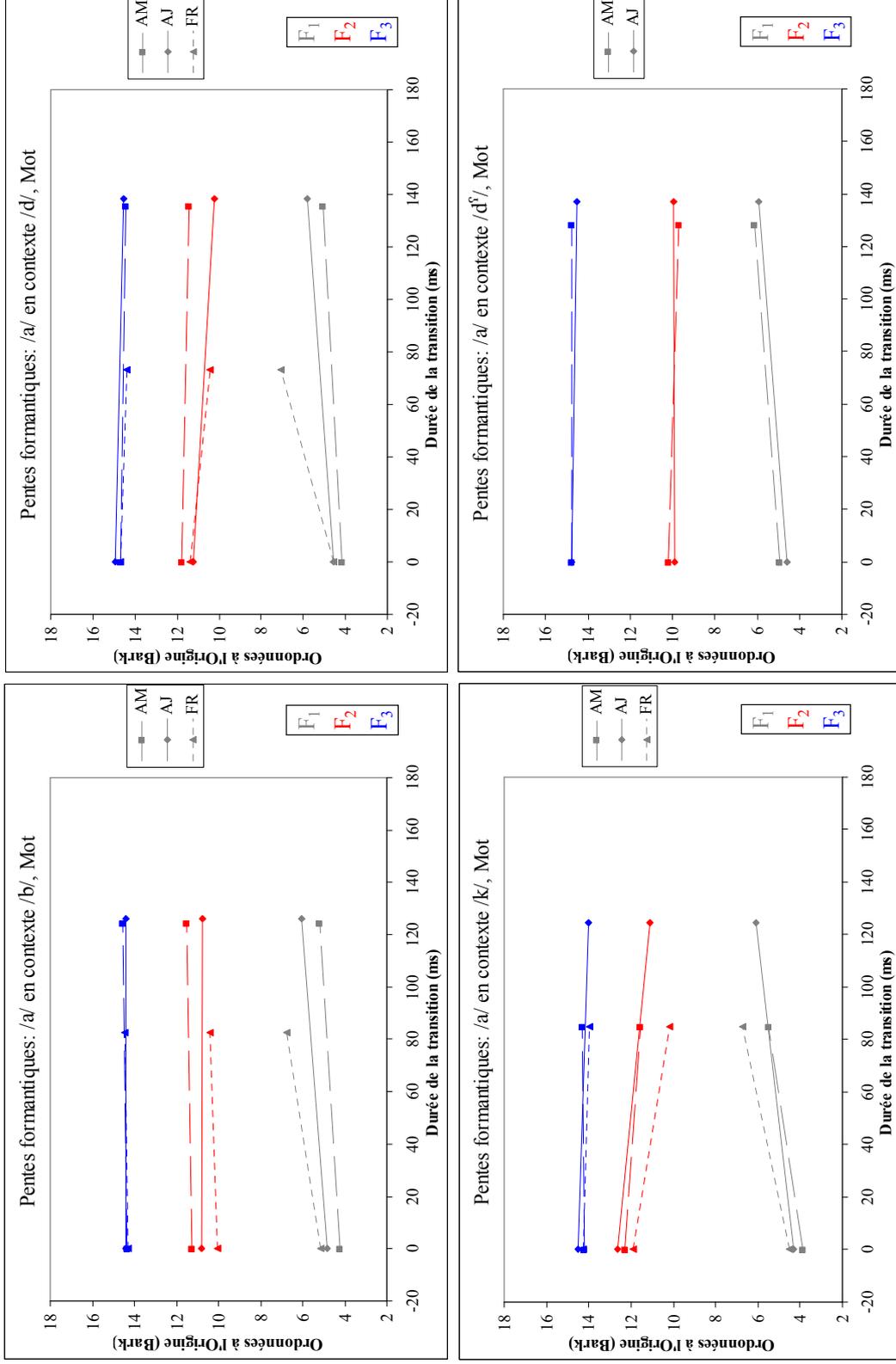


Figure 48 : Pentes formantiques de F₁, F₂ & F₃ sans normalisation du temps, pour la voyelle /a/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

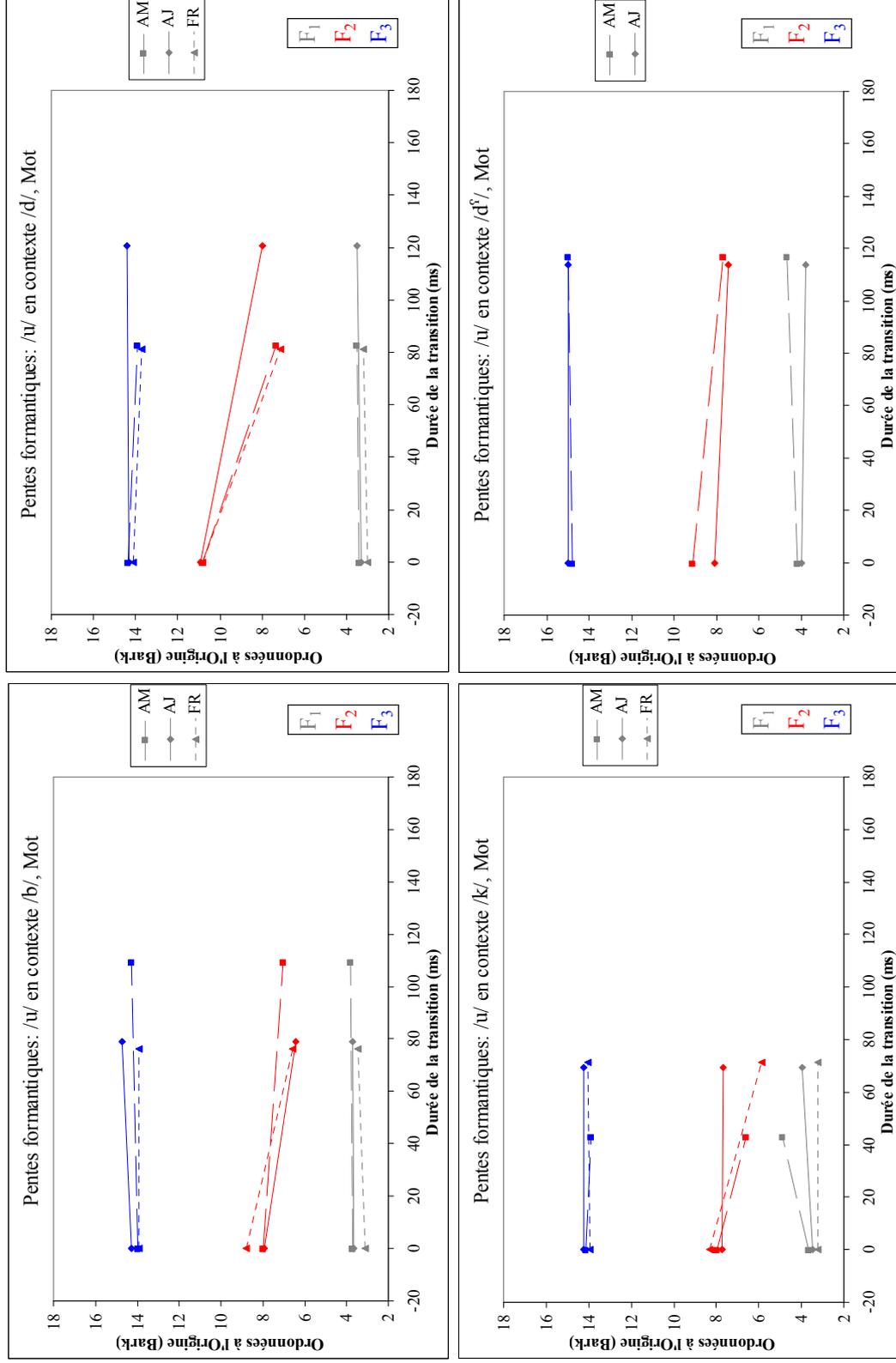


Figure 49 : Pentas formantiques de F₁, F₂ & F₃ sans normalisation du temps, pour la voyelle /u/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

Les différences d'inclinaison des **pent**es formantiques en fonction du lieu d'articulation suivent en partie les propositions de la théorie du « locus virtuel » ; i.e., des points différents en fonction des lieux bilabial, dental et vélaire. Néanmoins, nous observons des variations importantes liées à la qualité de la voyelle : quels que soient la langue et le lieu d'articulation, nous obtenons des valeurs de début de transitions plus élevées pour la voyelle /i/ et plus basses pour /u/ ($p < 0,001$). Nous obtenons pour la voyelle /u/ un seul « Locus virtuel » pour les lieux bilabial et vélaire ; la différence entre les deux lieux d'articulation s'observe à partir des pentes formantiques de F_3 . L'étude des influences du contexte /d^h/ montre des variations importantes liées à la « centralisation » de l'espace vocalique en AM par rapport à celui en AJ. Par souci de lisibilité de graphiques, nous avons décidé de ne présenter que les résultats moyens obtenus pour chaque langue et en fonction des différents lieux d'articulation. Nous nous sommes rendu compte que la variabilité inter-individuelle était beaucoup plus marquée en **dynamique** par rapport aux résultats des « cibles vocaliques » obtenues en **statique** (voir § 4.3). Nous avons présenté sur la Figure 50 les **pent**es formantiques de F_2 de la voyelle /i/ dans l'environnement /d/ produite en situation Mot en fonction des réalisations individuelles des locuteurs des trois langues.

La comparaison des **pent**es formantiques individuelles et moyennes (présentées sur la Figure 50) obtenues pour la voyelle /i/ dans l'environnement /d/ en situation de production Mot montre des différences importantes selon, d'une part, le point de départ des transitions formantiques de cette voyelle et d'autre part, le degré d'inclinaison de ses transitions. Sur la Figure 47, nous obtenons approximativement le même point de départ : 12,82 Bark en AM, 12,92 Bark en AJ et 12,62 Bark en FR ; avec des degrés d'inclinaison des pentes formantiques différentes (0,007 Bark/ms en AM, 0,01 Bark/ms en AJ et 0,004 Bark/ms en FR). Ainsi nous pouvons quantifier le changement formantique moyen de cette voyelle (obtenu à partir du point 0 jusqu'au milieu temporel), pour chaque langue dans cet environnement de : 0,87, 0,76, 0,38 Bark pour AJ, AM et FR respectivement (avec des durées moyennes de 87,8, 113,7 et 96,5 ms en AJ, AM et FR respectivement).

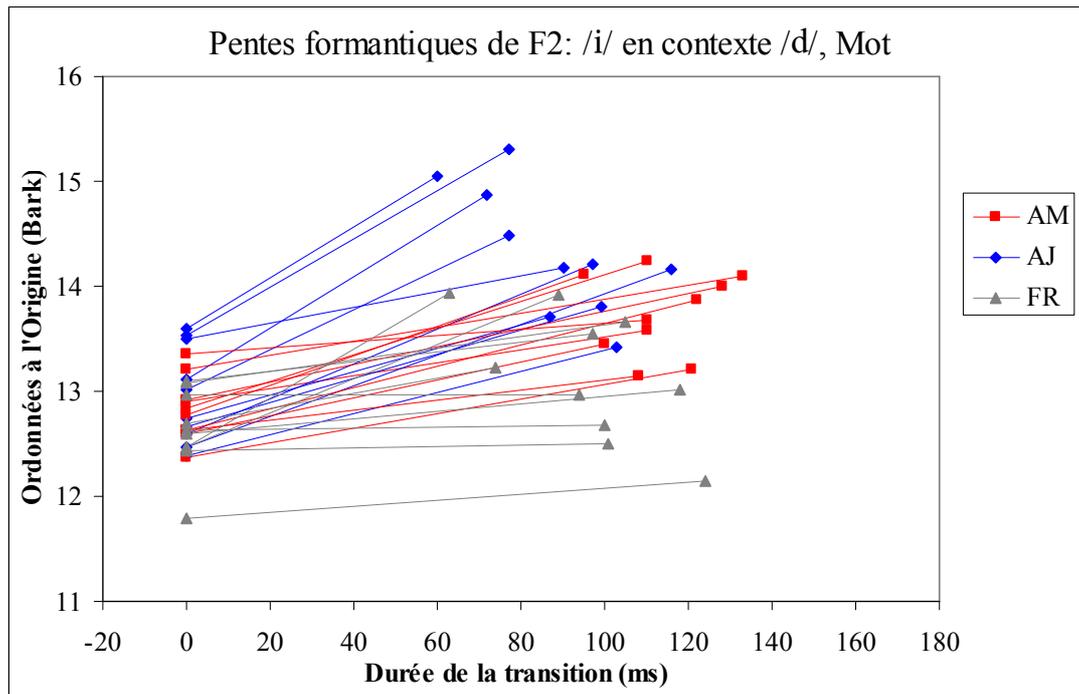


Figure 50 : Pentas formantiques individuelles moyennes, non normalisées, de F₂ de la voyelle /i/ produite dans la situation Mot dans l'environnement /d/ dans les trois langues.

Du point de vue des résultats individuels, nous observons des variations très marquées en fonction du point de départ des transitions formantiques ; du degré d'inclinaison et des changements formantiques obtenus. Ainsi, nous présentons dans le Tableau 12 ces différentes valeurs en fonction du locuteur qui présente le plus petit et le plus grand mouvement de formants dans les trois langues.

Tableau 12 : Valeurs moyennes des durées, des pentes formantiques, du début des transitions, des cibles vocaliques dérivées et des changements formantiques, sans normalisation, obtenues pour les locuteurs présentant les mouvements de formants les plus petits et les plus grands pour la voyelle /i/ en contexte Mot, dans l'environnement /d/, pour le formant F₂, dans les trois langues. Le numéro du locuteur dans chaque langue est entre parenthèses.

		durée	pente formantique	début transition	cible vocalique dérivée	changement formantique
AM	min (5)	110	0,002	13,35	13,68	0,26
	max (6)	110	0,01	12,83	14,24	1,12
AJ	min (10)	90	0,005	13,49	14,18	0,45
	max (3)	116	0,01	12,47	14,16	1,42
FR	min (2)	94	-0,00002	12,96	12,96	-0,002
	max (3)	89	0,01	12,48	13,91	0,93

Ces différences sont explicables à partir des valeurs de début des transitions formantiques et des cibles vocaliques dérivées, en fonction des locuteurs. À chaque fois que la

valeur de changement formantique est élevée, la valeur de début de transition est basse et la valeur de la cible vocalique dérivée est élevée. Ces variations de valeurs de pentes et de changements formantiques peuvent être corrélées soit à des différences anatomiques (morphologie du conduit vocal différente ; positions différentes des articulateurs, etc...) soit à des différences de stratégies individuelles. Dans tous les cas de figures, ces variations peuvent aider à la discrimination des locuteurs entre eux dans une structure syllabique, consonantique ou vocalique précise. Nous observons également des différences importantes liées en partie à la structure de la langue étudiée : pour le F_2 de la voyelle /i/, produite dans la situation de production Mot dans l'environnement consonantique /d/ (voir Figure 50 et Tableau 12), nous nous attendons à obtenir les **pentés formantiques** les plus importantes en FR et les plus faibles en AM, car, lorsque le F_2 de la voyelle /i/ est réduit, la **pente** obtenue est très « aplatie » par rapport à la **pente** « raide » obtenue pour une voyelle non réduite (voir Figure 46). Or, nos résultats montrent que les **pentés formantiques** obtenues en FR sont les plus « aplaties » et celles en AJ, les plus « raides ». C'est la modification de « cible vocalique » dérivée en FR qui a affecté le degré d'inclinaison de ses **pentés**. En effet, à partir des résultats obtenus en **statique**, nous avons montré que la voyelle /i/ était significativement plus « postérieure » (i.e., un F_2 plus bas) en FR par rapport à l'AM ou à l'AJ. Cette baisse de fréquence de F_2 est la raison de modification de degré d'inclinaison des **pentés**, car les trois langues présentent approximativement le même point de départ de la transition et la même durée de transition formantique.

Les différences de durées des transitions formantiques quantifiées dans le Tableau 11 sont observables directement sur les Figures 46 à 48. Nous obtenons des durées de transitions formantiques beaucoup plus courtes en FR par rapport à AM ou AJ. Nous pensons que cette différence peut affecter directement le degré d'inclinaison des transitions formantiques dans les trois langues. Afin de pallier ce biais, nous avons normalisé les durées de pentes en les situant dans un intervalle précis situé entre 0 et 0,5 (voir Analyse et traitement des données, § 4.3.2.1). Nous pensons que la normalisation du temps éliminera simplement les biais d'allongement des transitions et leur effet direct sur le degré d'inclinaison de celles-ci. Ainsi nous visualiserons directement les effets de « centralisation » des voyelles en AM par rapport à celles en AJ ou FR sur les **pentés formantiques** : des pentes « plates » indiquent une

« centralisation » des voyelles ; tandis que des pentes « raides » auront pour effet une « périphérisation » de l'espace vocalique. Dans la partie qui suit, nous présentons les résultats des **pentés formantiques** avec normalisation du temps.

4.3.2.2.2. Avec normalisation du temps

Les résultats graphiques des **pentés formantiques** obtenues avec normalisation du temps sont présentés sur les Figures 51 à 53 (uniquement en situation Mot, pour les deux autres situations de production, voir Annexes 4 & 5). Sur F_1 , les valeurs des pentés formantiques sont plus élevées pour la voyelle /a/ que pour les deux autres voyelles ($F(3, 4804)=349,35$; $p<0,001$) ; sur F_2 , en revanche, c'est la voyelle /u/ qui obtient les valeurs les plus élevées. L'explication la plus probable concerne les mouvements de la langue qui sont très variables pour la voyelle /a/ sur F_1 ; et l'avancement et le recul de la langue pour la voyelle /u/ sur F_2 . Lorsque nous nous intéressons de plus près aux effets observés sur ces mêmes voyelles en fonction du lieu d'articulation ; nous obtenons toujours les mêmes effets pour /a/ sur F_1 et quel que soit le lieu d'articulation ($F(6, 4804)=2,62$; $p<0,05$) ; tandis que pour la voyelle /u/ sur F_2 , ces effets sont identiques seulement pour les consonnes /b d d^f/ ($F(6, 4804)=28,76$; $p<0,001$) ; avec la consonne /k/, nous obtenons des **pentés formantiques** de F_2 pour la voyelle /u/ quasi nulles.

Les résultats graphiques montrent des différences non significatives des **pentés formantiques** obtenues en fonction des langues : nous obtenons des **pentés formantiques** homogènes pour les trois langues sur les deux axes F_1 ($p=0,27$) & F_2 ($p=0,08$) ; nous obtenons néanmoins des différences significatives liées à F_3 : c'est en AJ que les **pentés formantiques** sont les plus basses et en AM, les plus élevées ($F(3, 4804)=32,42$; $p<0,001$). Néanmoins, les résultats statistiques de l'interaction entre les langues, la situation de production, le lieu d'articulation et les voyelles montrent que, sur F_1 , c'est la voyelle /u/ produite dans l'environnement /b/ en situation de production Syllabe en AJ qui obtient la **penité formantique** la plus basse, tandis que c'est la voyelle /a/ produite dans l'environnement /k/ en situation de production Mot, en AJ, qui obtient la **penité formantique** la plus élevée ($F(24, 4804)=2,32$; $p<0,001$). Sur F_2 , et en valeur absolue, c'est la voyelle /i/ produite dans

l'environnement /b/ en situation de production Mot en FR qui obtient la **pente formantique** la plus basse, tandis que la voyelle /u/ produite dans l'environnement /d/ en situation de production Mot, en AJ, obtient la **pente formantique** la plus élevée ($F(24, 4804)=2,99$; $p<0,001$).

Nous n'observons pas d'homogénéité bien marquée entre les résultats obtenus avec ou sans normalisation du temps. Par exemple, la voyelle /i/ produite en contexte bilabial en FR obtient le même degré d'inclinaison en comparaison avec celles en AJ ou AM sans la normalisation du temps (voir Figure 47) ; tandis que cette même voyelle produite dans le même environnement consonantique avec la normalisation du temps obtient un degré d'inclinaison moins marqué en FR par rapport à l'AJ ou l'AM (voir Figure 51). Il semblerait que la durée de la transition formantique affecte son degré d'inclinaison dans les trois langues, mais cet effet ne change pas les résultats observés : le point de départ des transitions et la « cible vocalique » dérivée sont différents pour les trois voyelles dans les trois langues. Ainsi, nous obtenons des **pentés formantiques** moyennes par langue différentes en fonction de la voyelle, du formant ou du lieu d'articulation. L'effet de la « centralisation » de l'espace vocalique en AM quantifiée par le degré d'inclinaison des **pentés formantiques** moins marqué, semble être dépendante en grande partie de la durée vocalique. Toutefois, lorsque nous comparons entre les deux dialectes arabes, dans l'environnement /d^ʕ/, et plus spécifiquement sur F₂, nous observons une inclinaison des pentes formantiques moins importante pour les trois voyelles /i a u/ en AM par rapport à celle en AJ, indiquant ainsi une « centralisation » plus marquée en AM⁵⁵.

⁵⁵ Pour la voyelle /u/, nous prenons en compte la direction et la valeur de la **pente formantique** et observons ainsi une valeur de pente plus basse en AM (-1,56 en AJ et -3,45 en AM). Si l'on ne prend en compte que la valeur absolue, nous obtenons ainsi une inclinaison beaucoup plus forte en AM par rapport à AJ. Dans les deux cas de figure, nous observons un abaissement de la valeur de la cible vocalique en AM par rapport à celle en AJ, et ainsi une « centralisation » plus marquée en AM.

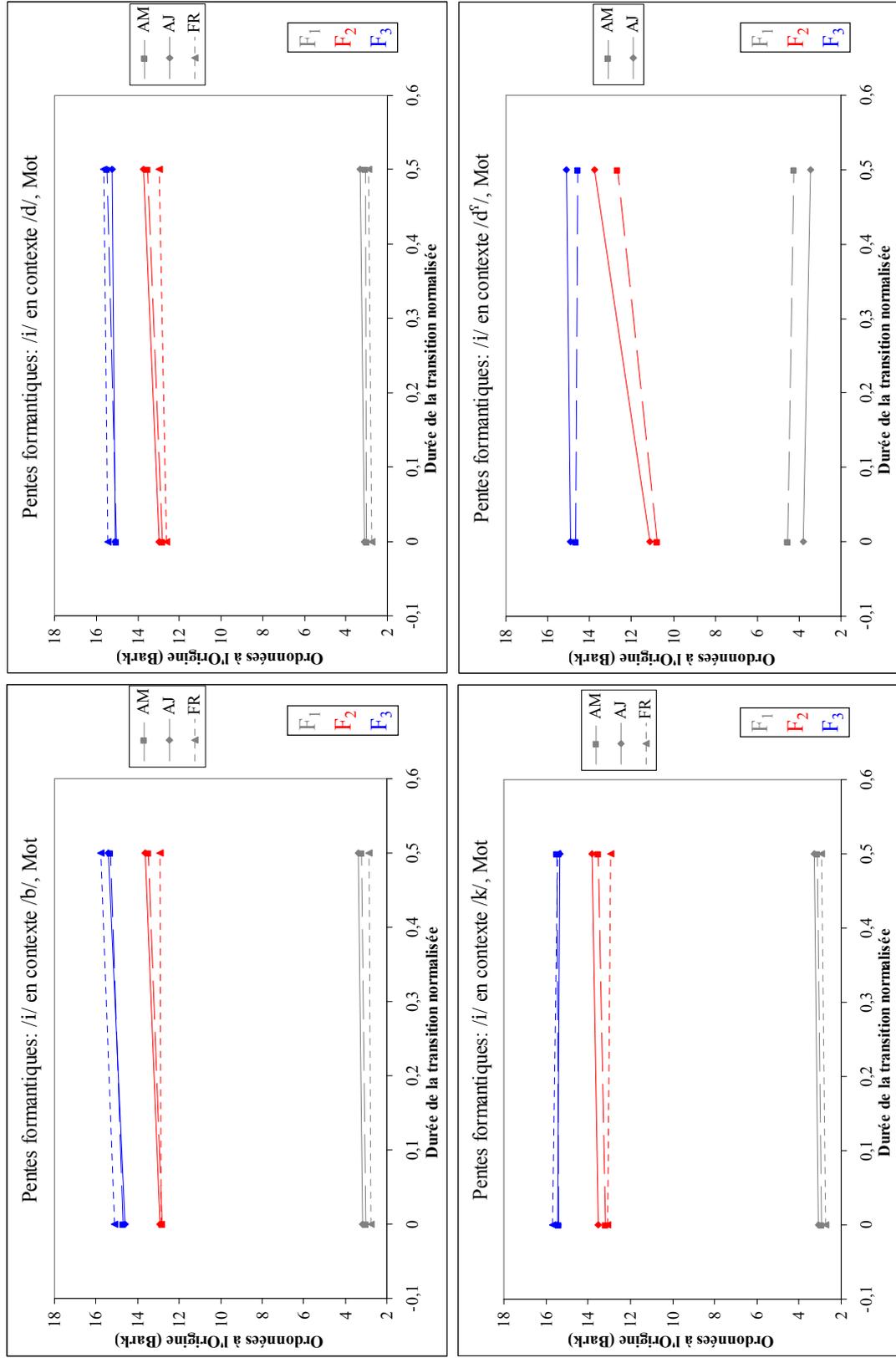


Figure 51 : Pentas formantiques de F_1 , F_2 & F_3 avec normalisation du temps, pour la voyelle /i/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

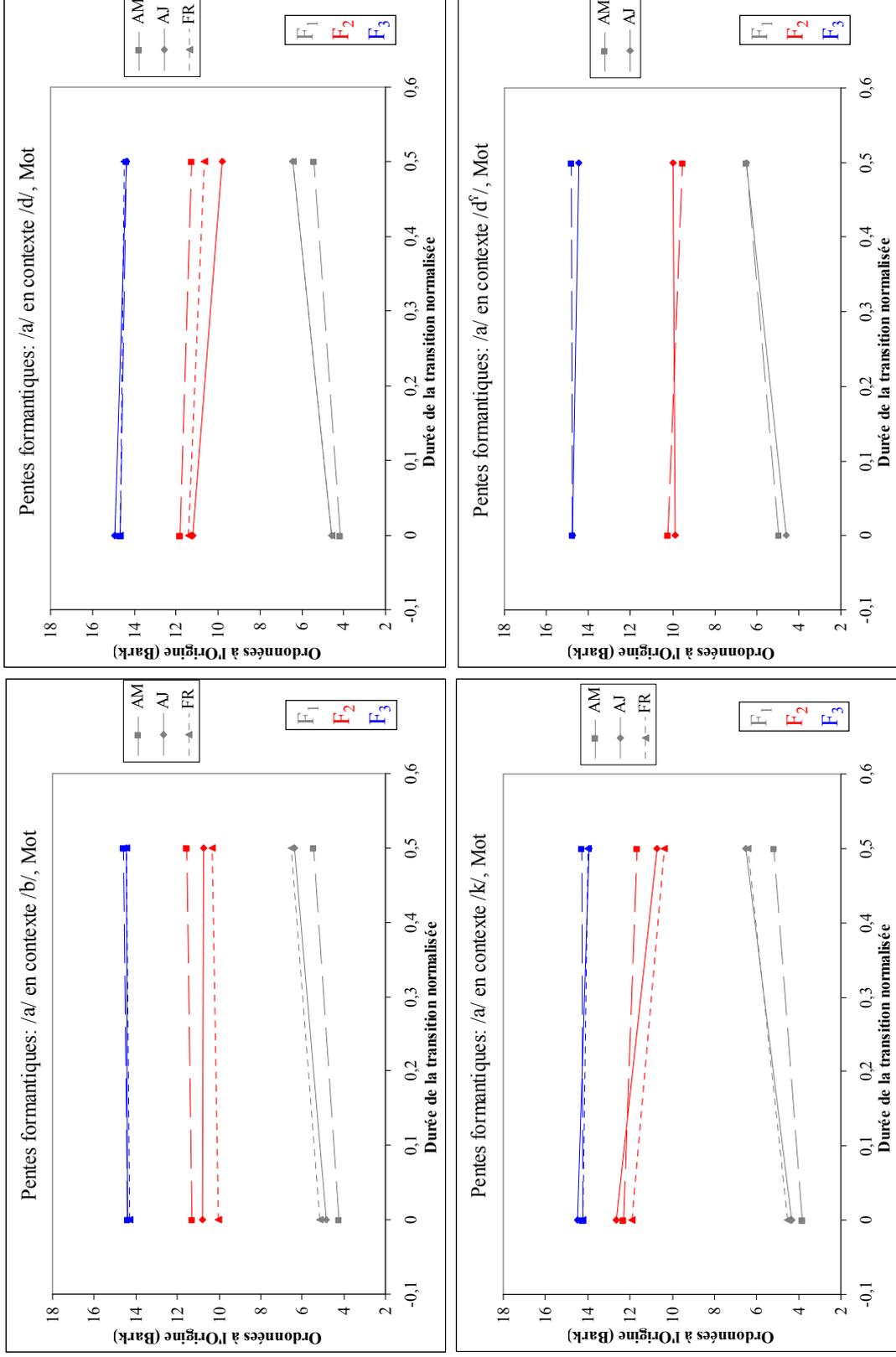


Figure 52 : Pentes formantiques de F_1 , F_2 & F_3 avec normalisation du temps, pour la voyelle /a/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

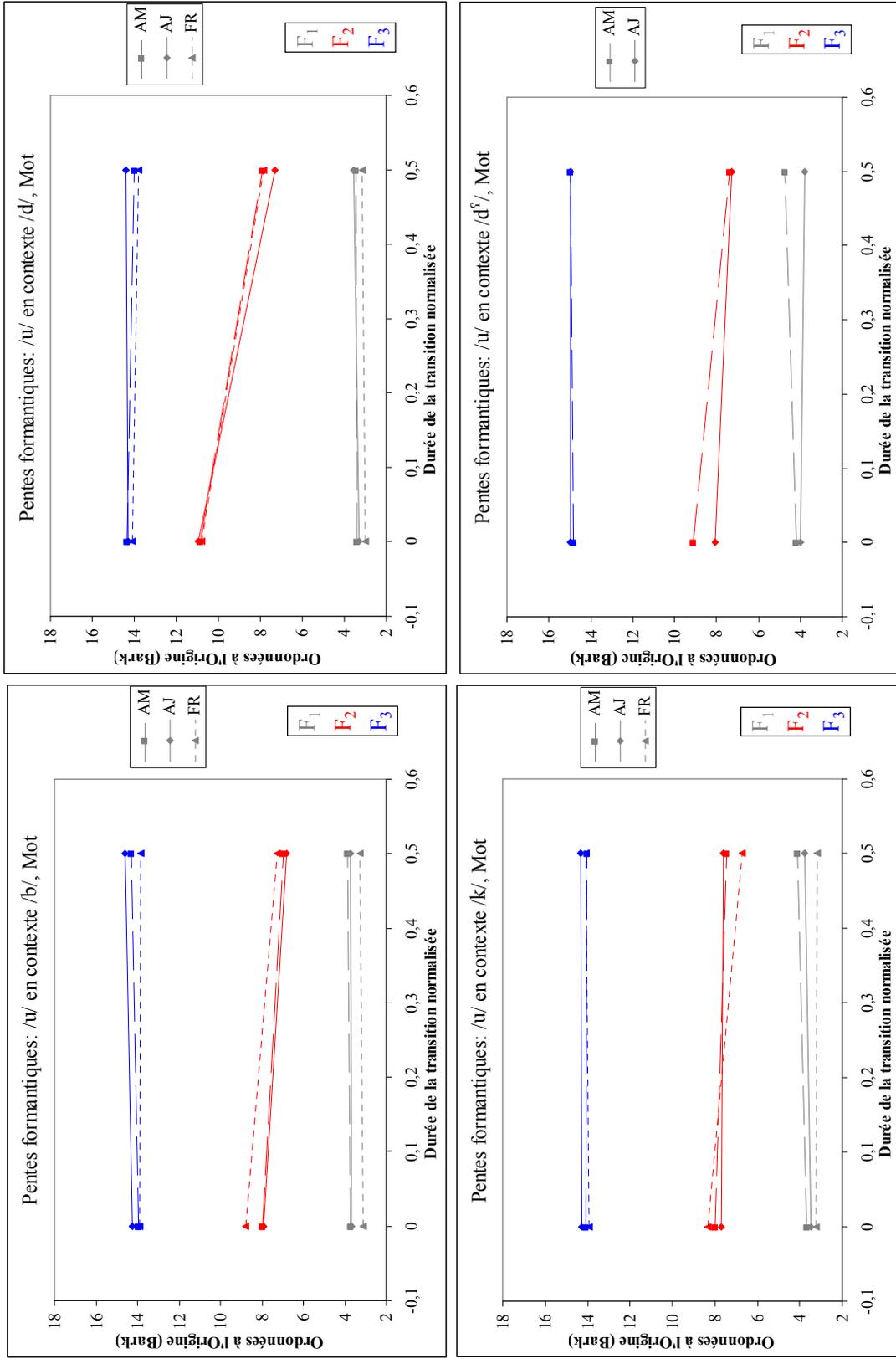


Figure 53 : Pentas formantiques de F₁, F₂ & F₃ avec normalisation du temps, pour la voyelle /u/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

Les résultats obtenus en **dynamique** vont dans le sens de ceux obtenus en **statique** et montrent l'existence de différences entre les deux dialectes arabes en fonction de cet environnement consonantique. Les effets de la densité des systèmes semblent affectés les **pentés formantiques** dans les trois langues en fonction du lieu d'articulation. Comme pour la comparaison sans normalisation de durée de la transition formantique, nous obtenons des différences importantes entre le point de départ des transitions, du degré d'inclinaison des **pentés formantiques** et de la cible vocalique dérivée ($p < 0,001$). Sur la Figure 51, les valeurs de début des transitions (12,82 Bark en AM, 12,96 Bark en AJ et 12,62 Bark en FR) et de la cible vocalique (13,54 Bark en AM, 13,73 Bark en AJ et 12,95 Bark en FR) sont approximativement homogènes ; celles de la **penne formantique** sont différentes (1,43 Bark/demie voyelle en AM, 1,55 Bark/demie voyelle en AJ et 0,65 Bark/demie voyelle en FR). Pour cette voyelle, le changement formantique moyen obtenu est de 0,77, 0,71, 0,33 Bark pour AJ, AM et FR respectivement.

Nous avons examiné les **pentés formantiques** obtenues dans les trois langues en fonction de chaque locuteur (voir la Figure 54 pour un exemple). Les résultats obtenus au niveau individuel montrent une homogénéité de la direction de la **penne formantique** de F_2 obtenue pour la voyelle /i/ produite dans l'environnement /d/ en situation de production Mot : toutes les **pentés formantiques** sont montantes de la consonne à la voyelle. Toutefois, nous observons un degré d'inclinaison de ces **pentés** différentes en fonction des locuteurs et des langues ($p < 0,001$). D'une façon générale, nous observons des **pentés formantiques** plus « raides » en AJ par rapport à celles en AM ; les pentes obtenues en FR sont globalement moins « raides ». Dans cet exemple, nous observons un effet inverse de ce à quoi nous nous attendions : avec l'effet de la densité des systèmes, la voyelle /i/ devrait être plus « centralisée » sur F_2 (i.e., une valeur formantique plus basse) en AM et plus « périphérique » en FR, or c'est en FR que la voyelle est plus « centralisée » sur F_2 .

C'est la valeur de $F_{2Milieu}$ qui semble être à l'origine de l'inversion des résultats. En effet, en AJ, la valeur $F_{2Milieu}$ est plus élevée, i.e., voyelle plus antérieure, que celle en AM ou en FR. Comme nous l'avons expliqué auparavant, la modification de la valeur de l'état sable affecte celle de l'onset vocalique et par conséquent l'inclinaison de la **penne formantique**

obtenue. Toutefois, ce résultat est consistant avec celui obtenu pour la même voyelle dans le même environnement consonantique en **statique** (voir Figure 34, p. 138) : l'effet de « périphérisation » en FR était observable sur F₁ plutôt que sur F₂.

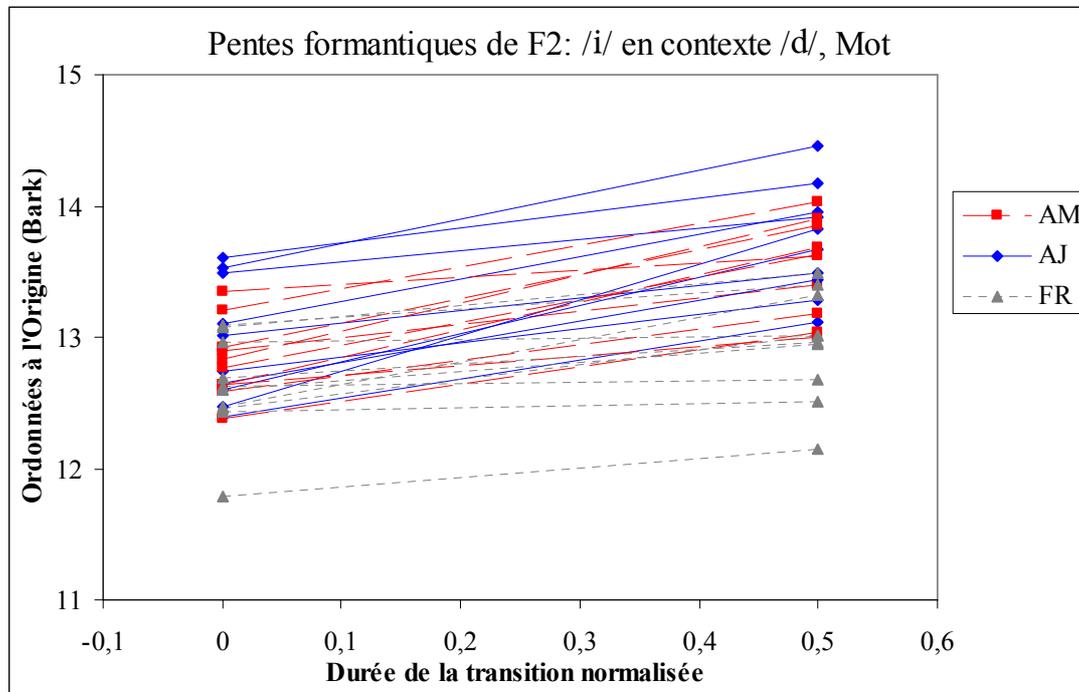


Figure 54 : Pentes formantiques individuelles moyennes, normalisées, de F₂ de la voyelle /i/ produite en situation de production Mot dans l'environnement /d/ dans les trois langues.

En examinant plus en détail les **pentés formantiques** obtenues pour chaque locuteur dans les trois langues, nous observons des degrés d'inclinaison variables. Ainsi, nous présentons dans le Tableau 13 les valeurs des **pentés formantiques**, du début des transitions, des cibles vocaliques dérivées et des changements formantiques obtenues pour les locuteurs ayant les mouvements de formants les plus petits et les plus grands.

Tableau 13 : Valeurs moyennes des pentes formantiques, du début des transitions, des cibles vocaliques dérivées et des changements formantiques, avec normalisation, obtenues pour les locuteurs présentant les mouvements de formants les plus petits et les plus grands pour la voyelle /i/ en contexte Mot, dans l'environnement /d/, pour F₂, dans les trois langues. Entre parenthèses sont marqués les numéros des locuteurs dans chaque langue.

		pente formantique	début transition	cible vocalique dérivée	changement formantique
AM	min (5)	0,519	13,35	13,61	0,26
	max (6)	2,15	12,83	13,90	1,07
AJ	min (10)	0,848	13,49	13,92	0,42
	max (3)	2,71	12,47	13,83	1,36
FR	min (2)	0,10809	12,96	13,02	0,054
	max (3)	1,69	12,48	13,32	0,85

4.3.3. Discussion des résultats inter-langues en statique vs en dynamique

Les résultats obtenus dans les parties précédentes montrent l'existence de différences significatives entre les voyelles /i a u/ produites dans les trois langues ; différences qui peuvent être imputées aux effets de la densité des systèmes vocaliques, car l'organisation des systèmes vocaliques dans les trois langues est différente.

En effet, en **statique**, nous avons montré que les espaces vocaliques, des trois voyelles /i a u/ et de toutes les voyelles, obtenus en FR sont significativement plus grands que ceux en AM ; en AJ, les espaces vocaliques se situent entre les deux langues. Plus spécifiquement, c'est sur l'axe F₁, que les espaces vocaliques du FR sont les plus grands par rapport aux deux autres langues. Ces résultats sont systématiques quels que soient la situation de production et le lieu d'articulation. La taille des aires de dispersion des espaces vocaliques est significativement plus grande en FR par rapport à l'AM ; aucune différence n'a été obtenue pour la comparaison entre les aires de dispersion en FR et en AJ. Ces résultats vont dans le sens des propositions de la **théorie de la dispersion adaptée** (Lindblom, 1986), selon laquelle la taille des espaces vocaliques est fonction de leur densité : plus une langue a de voyelles et plus la taille de l'espace vocalique est grande, pour permettre une meilleure discrimination entre les voyelles. Lorsque nous avons examiné la taille des ellipses de dispersion de chacune des voyelles étudiées (i.e., /i a u/), nos résultats ont montré l'effet inverse de celui prédit par la **théorie quantique** et la **théorie de la dispersion adaptée**. À

partir des prédictions avancées par ces deux théories, nous savons que la première explique que la densité des systèmes vocaliques n'affecte en rien la taille des espaces vocaliques ni l'organisation des voyelles dans l'espace acoustique. En ce qui concerne la dispersion intra-catégorie vocalique, la **théorie quantique** explique que puisque les trois voyelles /i a u/, ou les « *Hot Spots* » sont produites dans des zones de stabilité de l'espace phonétique, elles se situent aux « extrémités » de l'espace acoustique et par conséquent la taille de leurs ellipses ne dépendent pas de la densité des systèmes vocalique. À l'inverse, la **théorie de la dispersion adaptée** explique que comme la densité des systèmes vocaliques affecte l'organisation des voyelles à l'intérieur de l'espace vocalique, la taille de cet espace dépend donc de la distance **maximale** et **adaptée** entre les voyelles. Par conséquent, d'après cette théorie, la taille des ellipses de chaque voyelle dépend également de la densité des systèmes : plus une langue a des voyelles et plus la taille des ellipses de chaque voyelle sont petites afin de garantir le maximum de distinctivité entre elles, et *vice versa*. Nos résultats ont montré que la taille des ellipses des voyelles est indirectement corrélée à la densité des systèmes, car c'est en AM que les ellipses de chaque voyelle sont les plus petites et en AJ, les plus grandes ; le FR quant à lui, présente les ellipses de dispersion de chaque voyelle les plus petites. Ainsi, l'AM semble respecter les propositions décrites dans le Cas N°. 2 ; l'AJ et le FR, le Cas N°. 3 : un espace vocalique « très » réduit avec une dispersion intra-catégorie vocalique réduite (le cas de l'AM) ; et un espace vocalique plus grand avec une dispersion intra-catégorie vocalique grande (le cas de l'AJ et du FR).

En **dynamique**, nous avons montré que les **pentés formantiques**, obtenues à partir de l'analyse de régression linéaire⁵⁶, sont en moyenne, plus « raides » en FR par rapport aux deux autres langues ; l'AM quant à lui présentent les **pentés** les plus « aplaties », et plus particulièrement sur F_2 & F_3 ; aucun effet n'est observé sur F_1 . Afin de mettre en évidence les effets de la densité des systèmes vocaliques sur le degré d'inclinaison des **pentés formantiques**, nous avons expliqué que le degré d'inclinaison de la transition formantique, quantifiée de l'**onset** (temps à 0 ms) à la « cible vocalique » (temps à 50% de la durée vocalique), rendra compte davantage des effets de la densité des systèmes : pour une voyelle

⁵⁶ Mais également les analyses de régressions polynomiales de 2^{ème} et de 3^{ème} ordre.

précise, une **pente formantique** « forte » indique un degré de « centralisation » faible, car la réalisation de la voyelle se rapprocherait davantage de la production d'un « cible vocalique » *Hyper-Articulée*. Étant donnée que la valeur formantique obtenue au milieu temporel de la voyelle affecte celle de l'**onset** (voir explications dans le texte et sur la Figure 46, p. 164), la **pente formantique** obtenue sera plus « forte » lorsque la production de la voyelle est *Hyper-Articulée*.

Nous avons observé des différences significatives entre les trois langues liées au degré d'inclinaison des **pentés formantiques** ; au point de départ des transitions, i.e., l'**onset** vocalique et à la cible vocalique dérivée. Ces différences semblent être corrélées aux situations de production et aux lieux d'articulation. En fonction des situations de production, nous avons montré que les **pentés formantiques** des voyelles produites en situation Mot, sont beaucoup plus « raides » que celles en Syllabe ; quant aux **pentés** obtenues en Isolation, elles sont très « aplaties » mais « non-nulles ». Ceci indique que lorsque les voyelles sont produites en Isolation, elles ne correspondent pas à des « cibles vocaliques » **soutenues**. Nous avons effectué auparavant des calculs des CSIV (Changements Spectraux Inhérents aux Voyelles, Nearey & Assmann, 1986 ; Hillenbrand *et al.*, 1995, 2001 ; Nearey, 1989 ; entre autres) sur les voyelles produites dans les trois situations de production en extrayant deux valeurs fréquentielles à 20% et à 80%, ces deux positions correspondent à celles proposées dans la littérature, et ont été déterminées car elles se situent à l'endroit où les effets coarticulatoires des consonnes adjacentes sur les voyelles sont limités. Les résultats de cette étude ont montré que même en situation Isolation, les voyelles ne sont jamais « stable » de l'**onset** à l'**offset** vocalique ; elles ont plutôt des trajectoires formantiques propres à chacune. Bien que les résultats obtenus avec les CSIV soient intéressantes plus particulièrement en ce qui concerne les effets observés en Isolation ; nous avons décidé de ne pas les exploiter, car la comparaison entre les trois langues, les trois situations de production et les lieux d'articulation était impossible, à cause des structures syllabiques et consonantiques différentes des items utilisés. Nous exploiterons cette méthode d'analyse, bien que s'appuyant sur deux valeurs formantiques seulement, dans une nouvelle étude concernant les effets de la densité des systèmes sur les trajectoires formantiques propres aux voyelles, ou les **indices dynamiques intrinsèques**.

Lorsque le paramètre du temps, i.e., durée de transition formantique, est normalisé (en plaçant les valeurs formantiques dans un intervalle de 0 et 0,5, correspondant aux valeurs de l'**onset** et du milieu temporel, respectivement), les résultats obtenus ne sont pas en accord avec ceux obtenus sans la normalisation du temps. Ceci va à l'encontre de nos prédictions, car nous nous attendions à obtenir des **pentés formantiques** plus « raides » en FR et plus « aplaties » en AM, du fait de l'élimination de la durée de transition, qui à notre sens ne servait qu'à prolonger la transition formantique. Or, nous avons obtenu des **pentés formantiques** moyennes presque homogènes entre les trois langues (sur F_1 & F_2). La seule différence observée dans nos résultats concerne le point de départ des **pentés formantiques** et la « cible vocalique » dérivée. Nous pensons que même si le degré d'inclinaison des **pentés** n'est pas différent lorsque le temps est normalisé ; les différences de valeurs des points de départ de ces **pentés** et des « cibles vocaliques » dérivées peuvent être à l'origine de la discrimination entre les voyelles produites dans les trois langues étudiées et par conséquent sont dépendantes de la densité des systèmes.

Ces différents résultats obtenus en **statique** vs **dynamique** ont été évalués avec l'aide d'une analyse discriminante, qui a été réalisée avec une validation croisée. Nous avons décidé de présenter le résumé des principaux résultats des différentes analyses discriminantes par souci de lisibilité, le lecteur pourra consulter les taux de classification correcte en fonction de chaque trait utilisé (**statiques** ou **dynamiques**), dont les résultats sont présentés dans les parties qui suivent. Nous présentons sur la Figure 55 les taux de classification moyens des voyelles /i a u/ (voyelles longues en arabe) dans les trois langues produites dans les contextes /b d k/ dans les situations Mot, Syllabe et Isolation, en fonction des paramètres **statique** ou **dynamique**. Nous nommons paramètres **statiques**, toutes les analyses des formants des voyelles du point de vue entièrement **statique** (i.e., caractérisation des « cibles vocaliques ») ou **dynamique** avec la normalisation du temps (que nous présentons sur la Figure 55 en fond hachuré) sans inclure ni la durée vocalique ni la durée de transition. En ce qui concerne les paramètres **dynamiques**, nous les considérons soit à partir des analyses entièrement **dynamiques** des formants des voyelles (analyses de régressions linéaire et polynomiales de 2^{ème} et de 3^{ème} ordre, sans la normalisation du temps), soit à partir des analyses **statiques** avec le paramètre de la durée vocalique inclus.

L'examen des taux de classification moyens des analyses discriminantes, présentés sur la Figure 55, montre que l'utilisation de n'importe laquelle des analyses **dynamiques** améliore ces taux significativement. Ainsi, nous observons une augmentation moyenne significative de 11% en situation Mot et de 10% dans les deux situations Syllabe et Isolation lorsque les différents paramètres **dynamiques** sont utilisés dans l'analyse discriminante, (BinomialQ ; $p < 0,001$)⁵⁷. Cette augmentation confirme nos résultats, car nous nous attendions à obtenir des meilleur taux de classification en utilisant les différents paramètres **dynamiques**. L'examen attentif des résultats graphiques montre que même lorsque les analyses **dynamiques** sont utilisées avec le temps normalisé (analyses de régressions linéaire et polynomiales de 2^{ème} et de 3^{ème} ordre, avec normalisation du temps), le taux moyen de classification correcte augment significativement de 7% par rapport à l'utilisation des seules fréquences formantiques obtenues à l'état stable, ou les « cibles vocaliques ». Parmi les trois analyses **dynamiques** obtenues avec la normalisation du temps, l'analyse de régression de 3^{ème} ordre semble être celle qui caractérise le mieux les trajectoires formantiques, quelle que soit la situation de production, (avec un taux moyen non significatif supérieur de 2%, BinomialQ ; $p=1$).

En ce qui concerne les différents paramètres **dynamiques**, les résultats de l'analyse discriminante montrent que c'est l'analyse de régression polynomiale de 2^{ème} ordre qui semble caractériser le mieux les trajectoires formantiques, quelle que soit la situation de production (sans augmentation significatif des taux, BinomialQ ; $p=1$). L'utilisation du paramètre de la durée vocalique augmente significativement les taux de classification, plus particulièrement lorsque nous comparons les taux obtenus à partir des analyses **statiques** (avec les fréquences formantiques de l'état stable) en incluant et en excluant la durée

⁵⁷ Le BinomialQ (p, k, n) est une analyse statistique, proposée dans Praat, qui permet de calculer la probabilité pour que, dans un nombre précis d'expérience n , un événement avec une probabilité p intervienne avec un nombre de fois k . Dans notre étude, nous calculons ainsi la probabilité pour que le taux de classification obtenu en **statique** p , soit significativement différent du taux de classification en **dynamique**. La méthode de calcul utilisée est la suivante : BinomialQ (p, k, n) = BinomialQ (taux de classification en **statique**, taux de classification en **dynamique***la somme du nombre total de voyelles dans les deux conditions, la somme du nombre total de voyelles dans les deux conditions).

vocalique. Le taux moyen de la classification augmente de 13% en utilisant la durée vocalique (BinomialQ ; $p < 0,001$).

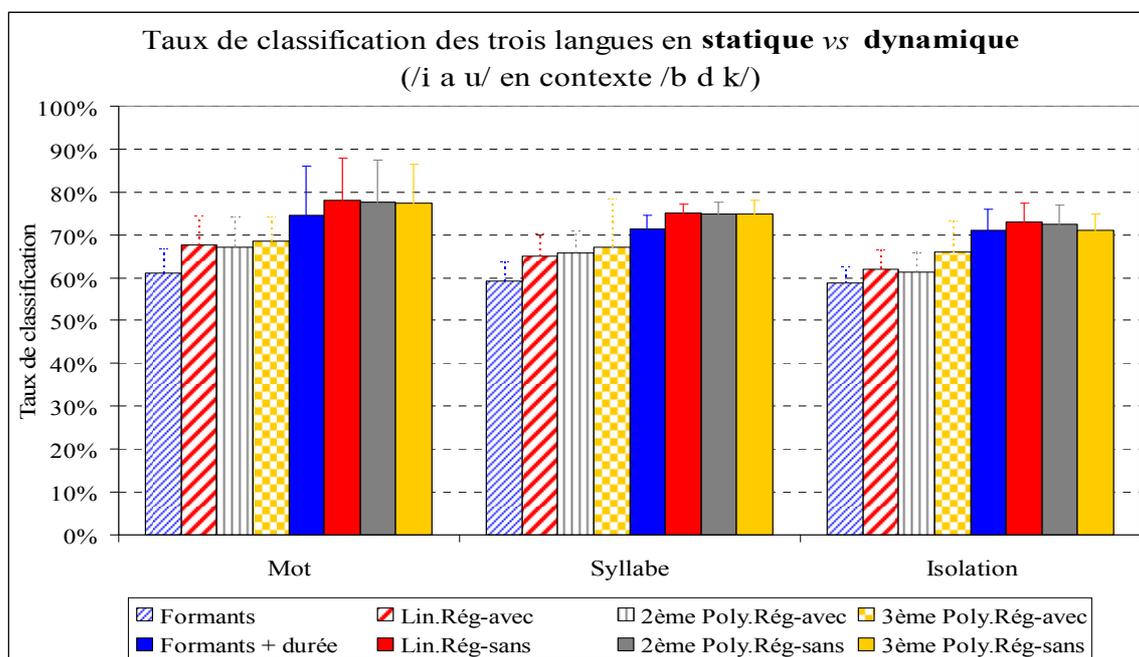


Figure 55 : Taux de classification correcte moyens des voyelles /i a u/ produites dans les contextes /b d k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur), dans les trois langues. Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} ou 3^{ème} ordre), sans = sans normalisation du temps, avec = avec normalisation du temps.

Nous présentons sur la Figure 56, les taux moyens de classification correcte de l'analyse discriminante en utilisant les différents paramètres **statiques** et **dynamiques** lors de la classification des trois voyelles /i a u:/ produites dans l'environnement /d^s/ dans les trois situations de production Mot, Syllabe et Isolation. Nous observons que l'analyse permet de discriminer l'arabe marocain de l'arabe jordanien dans cet environnement consonantique avec des taux moyens de 76% et de 79% en **statiques** et **dynamiques**, respectivement (la différence n'est pas significative, BinomialQ ; $p = 1$). Nous obtenons une augmentation significative de 10% du taux moyen de classification avec les différentes analyses **dynamiques** (régressions linéaire et polynomiales) obtenues avec la normalisation du temps en comparaison avec les taux obtenus avec l'analyse entièrement **statique** des voyelles (à partir des fréquences formantiques à l'état stable) (BinomialQ ; $p < 0,001$). Le même effet est obtenu lorsque les taux de classification moyens obtenus avec les différentes analyses **dynamiques** (régressions linéaire et polynomiales) sans la normalisation du temps sont

comparés à ceux obtenus avec l'analyse **statique** des voyelles (durée vocalique incluse), i.e., une augmentation significative des taux de 10% (BinomialQ ; $p < 0,001$). Nous n'obtenons pas d'effet systématique d'amélioration des taux de classification en augmentant l'ordre de l'analyse de régression, car tantôt c'est la régression polynomiale de 2^{ème} ordre qui améliore les taux (en situation Mot), tantôt c'est la régression linéaire (en situation Isolation).

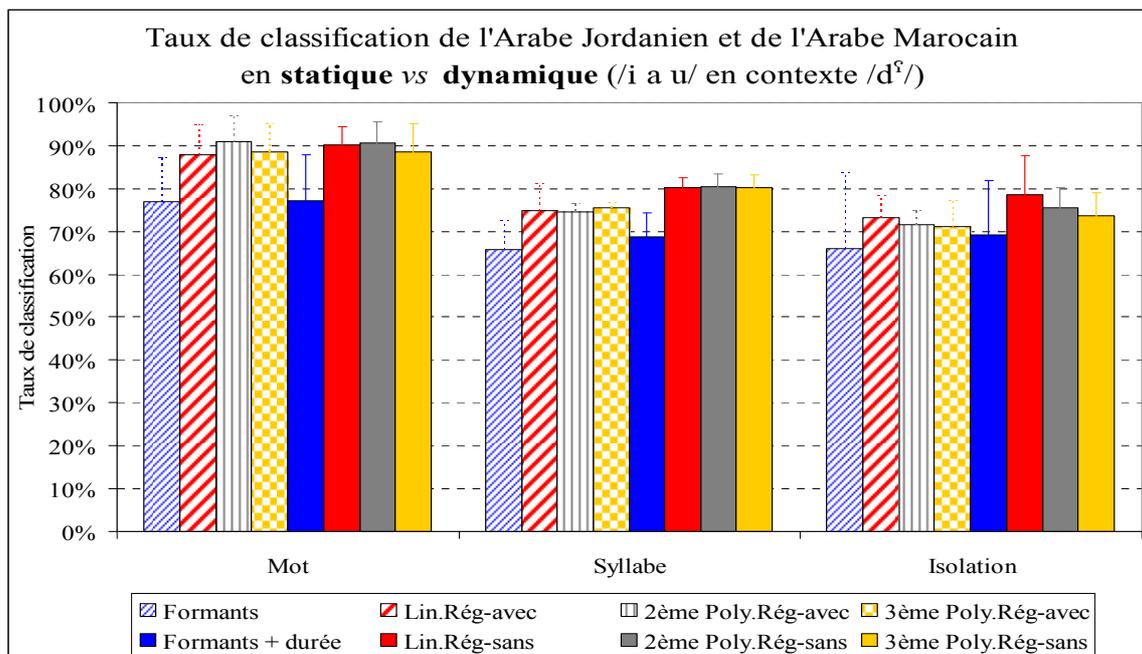


Figure 56 : Taux de classification correcte moyens des voyelles /i a u/ produites dans le contexte /d^ɣ/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur) en arabe jordanien et arabe marocain. Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} ou 3^{ème} ordre), sans = sans normalisation du temps, avec = avec normalisation du temps.

Nos résultats montrent que l'utilisation de la régression linéaire n'est pas suffisante pour décrire les trajectoires formantiques engendrées par l'environnement consonantique et/ou syllabique, en revanche, une analyse de régression polynomiale de 2^{ème} ou de 3^{ème} ordre apporte plus de précision quant à la quantification de ces trajectoires. Dans certains cas, il semblerait que l'utilisation de l'analyse de régression polynomiale de 3^{ème} ordre ne soit pas pertinente pour l'étude des voyelles monophthongues, sauf dans le cas où le temps est normalisé (voir Figure 55). Ces résultats sur la non pertinence de l'augmentation de l'ordre de l'analyse de régression polynomiale pour l'étude des voyelles monophthongues vont dans le sens des résultats obtenus par Kirsty McDougall (McDougall, 2006) qui expliquent en effet que l'analyse de régression polynomiale de 3^{ème} semble être le meilleur moyen d'obtenir une

description fidèle des trajectoires formantiques propres aux voyelles diphtongues en anglais britannique et australien. La régression polynomiale de 2^{ème} ordre semble être suffisante pour modéliser les trajectoires formantiques des monophthongues (McDougall, communication personnelle).

Ces résultats indiquent que, d'une part, les différentes informations **dynamiques** contenues dans la structure spectrale de la voyelle, qui sont la résultante soit des trajectoires **dynamiques** propres à la voyelle, ou les indices **intrinsèques**, soit des trajectoires **dynamiques** causées par l'entourage consonantique, ou les indices **extrinsèques**, sont d'une grande importance pour une meilleure caractérisation des voyelles /i a u/ quelles que soit la langue et la situation de production. D'autre part, ces résultats montrent que la simple utilisation du paramètre de la durée vocalique, en comparant les taux de classification avec les analyses **statiques** obtenus avec ou sans la durée, semble permettre à l'analyse discriminante de séparer les voyelles « correctement ». Nous supposons que les résultats obtenus avec les analyses discriminantes nous permettent d'appréhender le rôle des différentes informations **dynamiques** en production de la parole mais également en perception de la parole, car nous supposons que si ces informations sont pertinentes pour l'analyse discriminante, ceci indique qu'elles soient les plus saillantes pour discriminer entre les voyelles. Par conséquent, nous pouvons supposer qu'en perception, ces mêmes informations soient prises en compte par les auditeurs de chaque langue (voir Chapitre 5). Dans la partie qui suit, nous présentons en détail les résultats des différentes analyses discriminantes que nous appliquées à nos données en **statique** et en **dynamique**.

4.3.4. Résultats de l'analyse discriminante⁵⁸

Dans cette partie, nous nous intéressons aux résultats des différentes analyses discriminantes que nous avons appliquées sur nos données. Le but de cette analyse est avant tout une validation de nos résultats et plus particulièrement l'évaluation du rôle des paramètres **dynamiques** en comparaison avec les **statiques** dans l'analyse des voyelles dans les trois langues. Nous évaluerons également les différents paramètres **dynamiques** que nous avons utilisés dans ce travail, à savoir, les coefficients de la régression linéaire et les régressions polynomiales de 2^{ème} et 3^{ème} ordre.

4.3.4.1. En statique

Nous avons appliqué une analyse discriminante avec validation croisée⁵⁹ ayant comme paramètres les valeurs des trois premiers formants des voyelles /i a u/ en incluant et en excluant la durée vocalique. Les taux de classification correcte des trois langues en fonction de l'interaction entre les situations de production, les lieux d'articulation et les voyelles sont présentés dans le Tableau 14.

⁵⁸ Les taux de classification correcte de l'analyse discriminante sont variables et quelques fois, ils sont faibles. L'utilisation que nous faisons de ces taux n'est que comparative avec les taux de classification avec les indices dynamiques. Le taux de classification de l'analyse discriminante est considéré comme significatif si l'analyse statistique Box's M, proposée dans l'Analyse donne des résultats significatifs. Le Box's M est une analyse statistique qui propose de valider (ou non) l'hypothèse nulle selon laquelle la variance/covariance pour la totalité de l'échantillon est identique en fonction de toutes les données. Autrement dit, s'il existe des différences entre les sous-groupes testés, la moyenne et la variance vont être significativement différentes. La valeur de Box's M doit être la plus élevée possible : plus la valeur de Box's M est élevée, plus le seuil de significativité est bas (i.e., la probabilité est inférieure à 0,001). Si le seuil de significativité est supérieur à 0,05, l'analyse discriminante n'est pas valide (i.e., on ne peut pas effectuer une classification des données). Dans ce cas, nous mentionnerons ces taux et tenterons de les expliquer.

⁵⁹ La validation croisée est une méthode d'estimation des performances d'un modèle de classification à partir d'objets n'ayant pas servi à la conception du modèle. Dans notre étude, nous utilisons la méthode *leave-one-out* qui effectue la classification d'un objet précis à partir des données d'apprentissage, constituées de tous les objets moins l'objet à classifier.

Tableau 14 : Taux de classification correcte des trois langues en **statique** en fonction de la situation de production, du lieu d'articulation (/b d k/ entre les trois langues et /d^ʕ/ entre les deux dialectes arabes) et des voyelles /i a u/. La valeur entre parenthèses indique le taux de classification correcte sans durée de vocalique. (Taux significatifs, Box's M ; p<0,001)

		/b/	/d/	/k/	/d ^ʕ /
Mot	/i/	76,0% (63,3%)	74,0% (66,7%)	62,0% (56,0%)	78,6% (80,6%)
	/a/	86,7% (61,3%)	91,9% (67,6%)	86,5% (66,9%)	66,0% (65,0%)
	/u/	63,5% (50,0%)	67,8% (59,1%)	63,9% (57,8%)	87,0% (85,0%)
Syllabe	/i/	74,7% (58,0%)	72,7% (67,3%)	67,3% (54,7%)	64,3% (58,2%)
	/a/	74,7% (60,0%)	73,6% (52,0%)	70,9% (63,5%)	67,0% (67,0%)
	/u/	66,2% (55,4%)	73,2% (61,7%)	69,6% (60,1%)	75,0% (72,0%)
Isolation	/i/	68,0% (58,7%)	73,3% (58,0%)	67,8% (59,1%)	58,2% (48,0%)
	/a/	76,7% (64,7%)	78,4% (61,5%)	75,0% (60,8%)	83,0% (84,0%)
	/u/	66,5% (54,7%)	69,1% (60,4%)	64,6% (51,0%)	66,7% (65,7%)

Les résultats de l'analyse discriminante, par langue en fonction de la situation de production, du lieu d'articulation et des voyelles /i a u/ (présentés dans le Tableau 14), montre la possibilité de différencier les trois langues avec des taux de classification élevés et ce en utilisant parfois une seule voyelle (voir l'exemple de la voyelle /a/ en contexte /d/ en Mot avec un taux de 91,9% en utilisant la durée). Le taux de classification correcte entre les trois langues est augmenté significativement en moyenne de 12% en utilisant la durée vocalique (72,39% avec la durée vocalique contre 59,64% sans durée), (BinomialQ ; p<0,001).

La classification de l'arabe marocain et de l'arabe jordanien en fonction des trois voyelles /i: a: u:/ produites dans l'environnement consonantique /d^ʕ/ a été possible avec un taux moyen significatif de 69,50% sans la durée vocalique et de 71,76% avec la durée vocalique (différence significative, BinomialQ, p<0,05)

Bien que les résultats de cette analyse aient dans le sens des différences observées dans les parties précédentes (en **statique**, voir § 4.3.1) ; les taux obtenus sont relativement bas. Nous nous sommes rendu compte que les confusions dans les classifications étaient à l'origine de la baisse des taux. Globalement, le FR était classifié presque tout le temps avec des taux avoisinant les 90% ; les deux dialectes arabes quant à eux ont été soit confondus entre eux, soit l'AJ avec le FR. Nous examinons ainsi ces confusions sur les Figures 57 et 58 en fonction de chacune des trois voyelles /i a u/ et le lieu d'articulation, en situation de

production Mot (pour les résultats graphiques des confusions obtenues en Syllabe et en Isolation, voir Annexes 6 et 7). Globalement, les voyelles du FR sont les mieux classifiées, celles en AJ, le moins. C'est la voyelle /a/ qui semble être la mieux classifiée parmi les trois voyelles. Nous observons également que la plupart des confusions dans la classification sont dues à la proximité des trois voyelles /i a u/ en AM et AJ ; celles du FR sont très peu confondues. Ceci va dans le sens des résultats statistiques obtenus, où les différences entre les voyelles n'étaient pas significatives entre l'AM et l'AJ.

L'utilisation de la durée vocalique comme paramètre classificatoire semble réduire les confusions dans la classification des trois voyelles /i a u/ dans les trois langues (voir Figures 57 et 58). Lorsque le paramètre de la durée vocalique n'est pas inclus dans l'analyse discriminante, nous obtenons moins de précision dans la classification (une moyenne de 61,7% sans durée contre 71,9% avec la durée vocalique). Nous observons approximativement les mêmes confusions avec les deux autres situations de production. Plus particulièrement, les voyelles de l'AM ont été la plupart du temps mal classifiées comme appartenant à l'AJ ; celles en FR ont été bien classifiées. Nous observons néanmoins qu'en Isolation, les voyelles /i a u/ ont été mal classifiées comme appartenant à l'AJ (20 à 30% des cas, voir Annexes 6 et 7). Les confusions dans la classification des voyelles /i a u/ produites dans l'environnement /d^h/ sont très importantes car le taux de classification moyen est proche des 70%. Ces confusions montrent que les valeurs formantiques des « cibles vocaliques » obtenues dans cet environnement sont très proches entre les deux dialectes arabes, i.e., les effets d'aperture et d'antériorisation observés sont comparables.

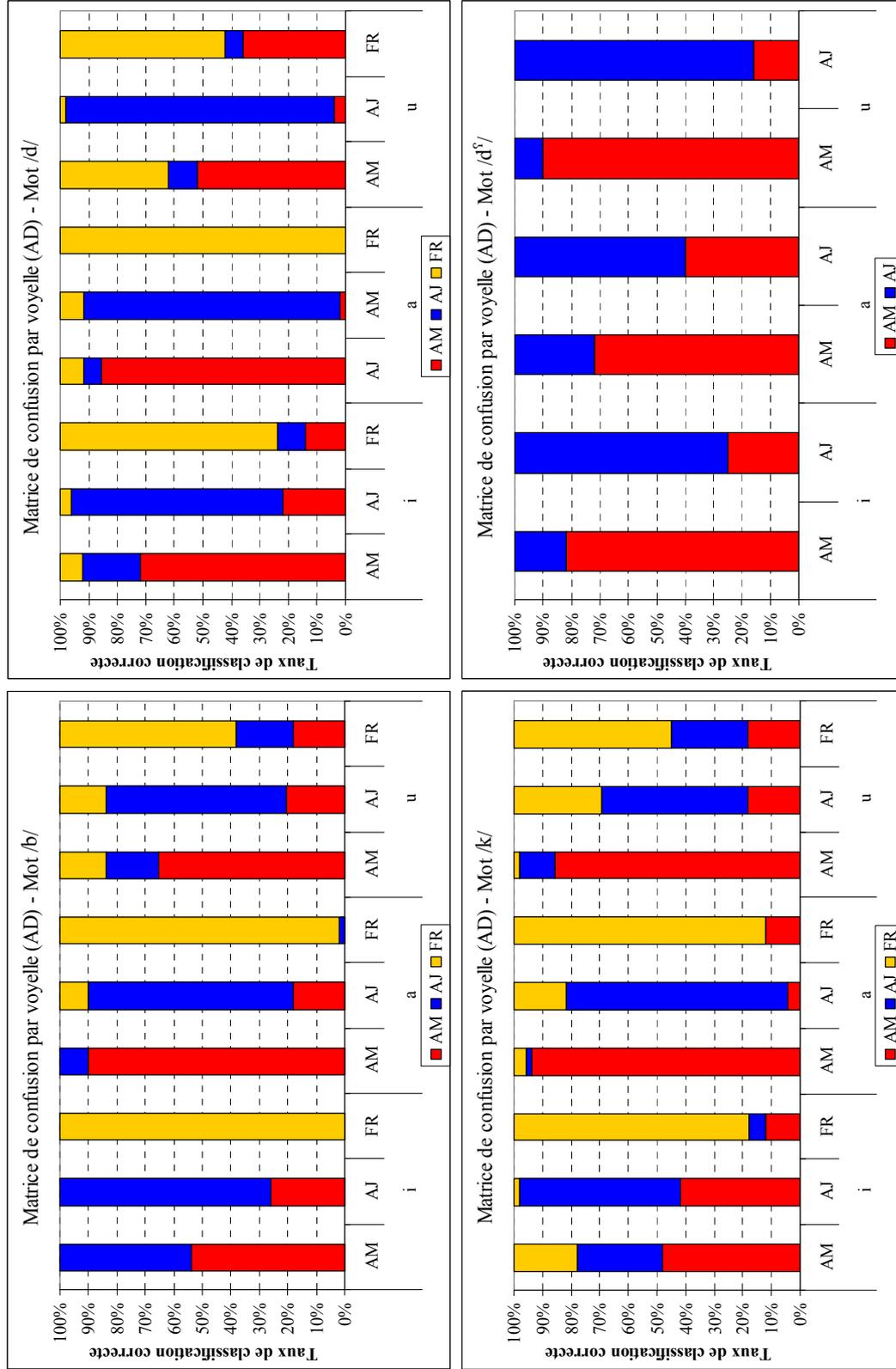


Figure 57 : Matrices de confusion de classification de chaque voyelle /i a u/ en **statique** avec la durée de la transition (AD) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d d^f/ k/ (/d^f/ entre AM & AJ).

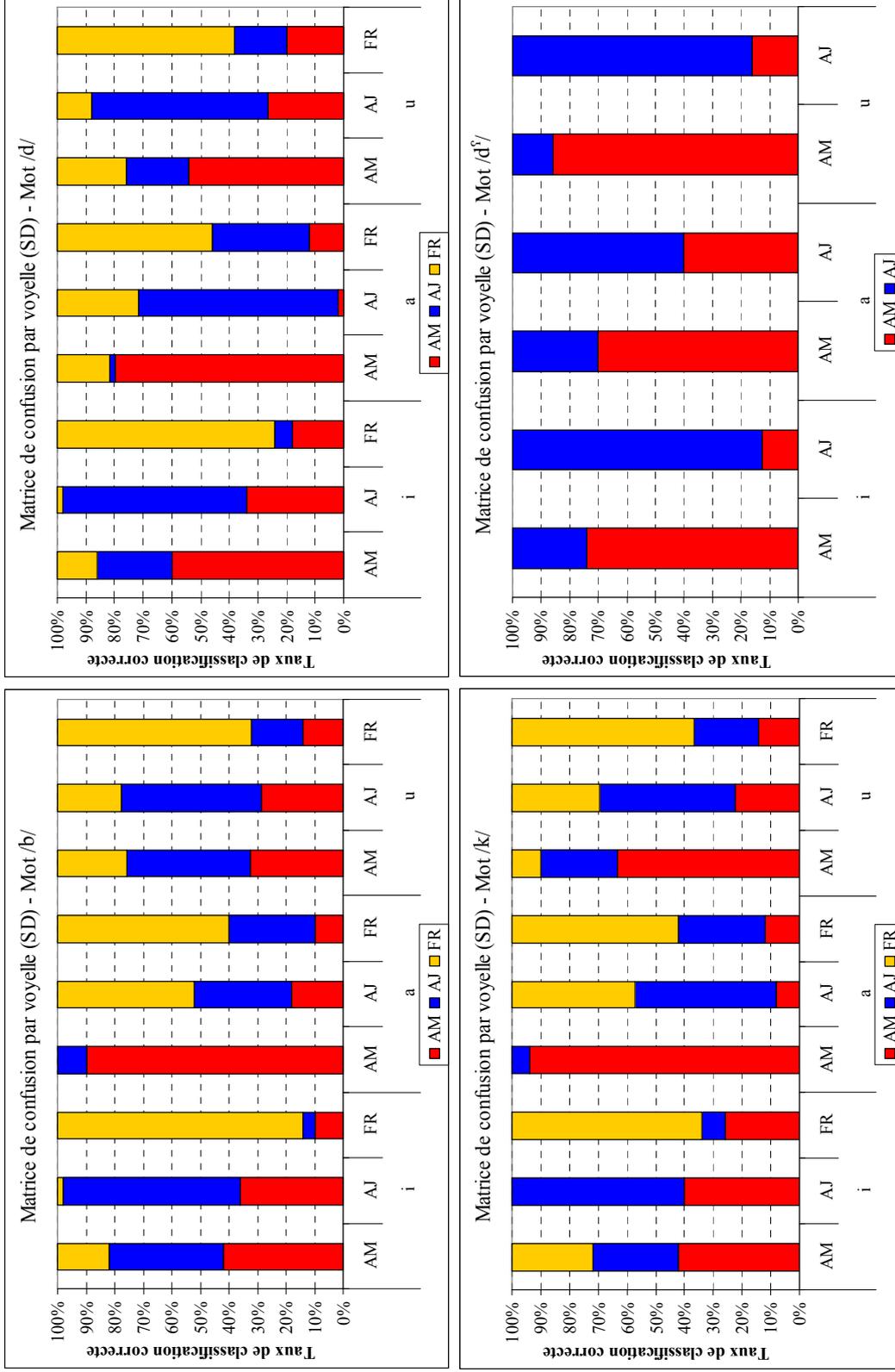


Figure 58 : Matrices de confusion de classification de chaque voyelle /i a u/ en **statique** sans la durée de la transition (SD) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f/ entre AM & AJ).

Une correspondance entre les résultats de l'analyse discriminante et nos résultats statistiques peut être établie. En effet, nous avons observé peu de différences significatives entre les voyelles /i a u/ de l'AM et de l'AJ, c'est pourquoi nous obtenons des confusions importantes entre les voyelles de ces deux dialectes ; néanmoins, la « centralisation » des trois voyelles observée en AM a été bien quantifiée par l'analyse discriminante, car plus de 60 à 70% des voyelles de l'AM ont été bien classifiées. En ce qui concerne le rapprochement entre les voyelles de l'AJ et du FR, l'analyse discriminante a permis d'en rendre, car souvent les confusions concernaient ces deux langues. La « périphérisation » de l'espace vocalique observée en FR a permis la classification de ses voyelles avec des taux supérieurs à 90%.

Les résultats de l'analyse discriminante que nous venons de présenter confirment les différences observées entre les trois langues, à savoir des voyelles « centralisées » en AM et « périphériques » en FR. Dans la partie qui suit, nous nous intéressons aux résultats de l'analyse discriminante en **dynamique** : soit les **pentés formantiques** obtenues à partir de la régression linéaire et polynomiales de 2^{ème} et 3^{ème} ordre. Nous évaluerons ainsi le rôle de ces paramètres dans la classification des voyelles des trois langues.

4.3.4.2. En dynamique

Nous présentons ici deux types d'analyses discriminantes. La première sera conduite à partir des coefficients obtenus avec la régression linéaire ; la seconde, avec ceux des deux régressions polynomiales de 2^{ème} et 3^{ème} ordre. Le but étant de voir quel ordre nous permettra de mieux discriminer les voyelles des trois langues.

4.3.4.2.1. Régression linéaire

Les paramètres utilisés dans cette analyse discriminante sont les valeurs des **pentés formantiques** et des ordonnées à l'origine F_1 , F_2 & F_3 , avec et sans normalisation du temps et en ajoutant la durée de la transition formantique lorsque le facteur temps n'est pas normalisé. Nous avons observé qu'à chaque fois que le paramètre de la durée de la transition est présent, il est utilisé par l'analyse discriminante comme le premier facteur permettant de discriminer entre les trois langues ; lorsque la durée est normalisée, c'est soit le paramètre de **pentés**

formantiques de F_2 soit l'ordonnée à l'origine de F_1 qui sont utilisés comme premiers paramètres permettant la séparation entre les trois langues. Les différences existantes entre les trois langues en ce qui concerne le degré d'inclinaison, le point de départ des transitions et des cibles vocaliques dérivées ont permis à l'analyse discriminante de séparer les voyelles dans les trois langues et dont les résultats sont présentés dans le Tableau 15.

Tableau 15 : Taux de classification correcte des trois langues à partir de l'analyse de régression linéaire, avec et sans normalisation du temps, en fonction de la situation de production, du lieu d'articulation (/b d k/ entre les trois langues et /d^s/ entre les deux dialectes arabes) et des voyelles /i a u/. Les valeurs entre parenthèses indiquent les taux avec normalisation du temps (Taux significatifs, Box's M ; $p < 0,001$).

		/b/	/d/	/k/	/d ^s /
Mot	/i/	76,7% (62,0%)	72,7% (66,7%)	66,7% (59,3%)	90,8% (89,8%)
	/a/	88,7% (78,0%)	91,2% (77,0%)	91,9% (74,3%)	86,0% (80,0%)
	/u/	72,3% (60,8%)	69,1% (65,1%)	72,8% (65,3%)	94,0% (94,0%)
Syllabe	/i/	76,0% (68,7%)	72,7% (68,0%)	73,3% (60,0%)	77,6% (69,4%)
	/a/	76,7% (57,3%)	76,4% (56,8%)	77,0% (68,9%)	82,0% (82,0%)
	/u/	77,7% (68,2%)	74,5% (69,8%)	71,6% (68,2%)	81,0% (73,0%)
Isolation	/i/	66,0% (56,0%)	71,3% (56,7%)	72,5% (63,8%)	68,4% (67,3%)
	/a/	76,0% (62,0%)	79,7% (68,9%)	75,7% (62,8%)	84,8% (77,8%)
	/u/	76,4% (62,8%)	69,1% (57,7%)	70,7% (68,0%)	82,8% (74,7%)

Ces résultats montrent la possibilité de discriminer les trois langues en fonction d'un lieu d'articulation et/ou d'une situation de production précise avec les valeurs des **pentés formantiques** et des ordonnées à l'origine des trois premiers formants des trois voyelles /i a u/, ainsi que la durée de la transition (uniquement lorsque la durée n'est pas normalisée). La classification correcte des langues par voyelle baisse significativement, en moyenne de 10,46% lorsque le temps est normalisé 64,93% avec normalisation et 75,39% sans normalisation), (BinomialQ ; $p < 0,001$). Il est normal d'observer cette baisse des taux de classification en normalisant le temps car comme nous l'avons expliqué, sans normalisation, la durée de la transition est le paramètre qui est utilisée toujours par l'analyse discriminante comme premier facteur aidant à la discrimination des voyelles entre les trois langues.

Les taux moyens de classification de l'analyse discriminante vont dans le sens des différences observées pour les **pentés formantiques** dans les trois langues. Dans la situation de production Mot, les **pentés formantiques** obtenues en FR ont été bien classifiées avec des

taux proches des 100% ; celles obtenues en AM ou en AJ ont été quelques fois confondues entre elles. Les résultats graphiques des confusions dans la classification des analyses discriminantes en situation Mot sont présentés sur les Figures 59 et 60 (pour la présentation graphique des confusions obtenues en Syllabe et en Isolation, voir Annexes 8 et 9). Comme avec les taux de confusions en **statique**, l'utilisation de la normalisation du temps baisse les taux moyens de classification et augmente les confusions surtout entre l'AM et l'AJ. Les taux de confusion dans la classification des voyelles dans les trois langues en situations Syllabe et Isolation montrent globalement les mêmes effets ; i.e., une meilleure classification des voyelles en FR, et une importante confusion dans la classification des voyelles de l'AM et de l'AJ. Le paramètre de la durée de la transition semble baissé les confusions dans la classification en situations Syllabe et Isolation. La classification des trois voyelles /i a u/ dans l'environnement /d^h/ a été améliorée en **dynamique** par rapport à celle en **statique** : le taux moyen de classification est proche des 85% avec des confusion moins importantes (même avec la normalisation du temps).

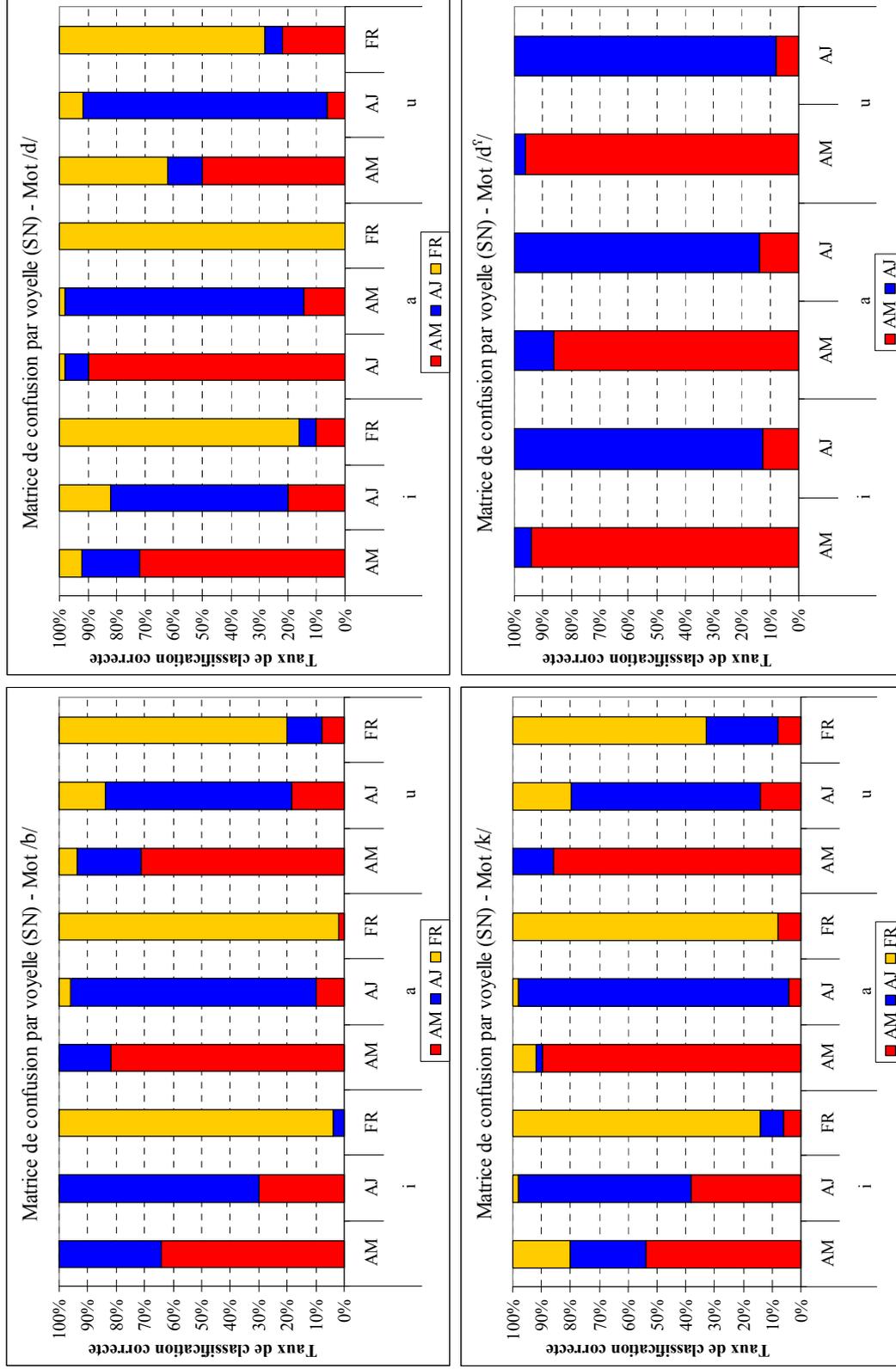


Figure 59 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression linéaire) sans normalisation du temps (SN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d d^h k/ (/d^h/ entre AM & AJ).

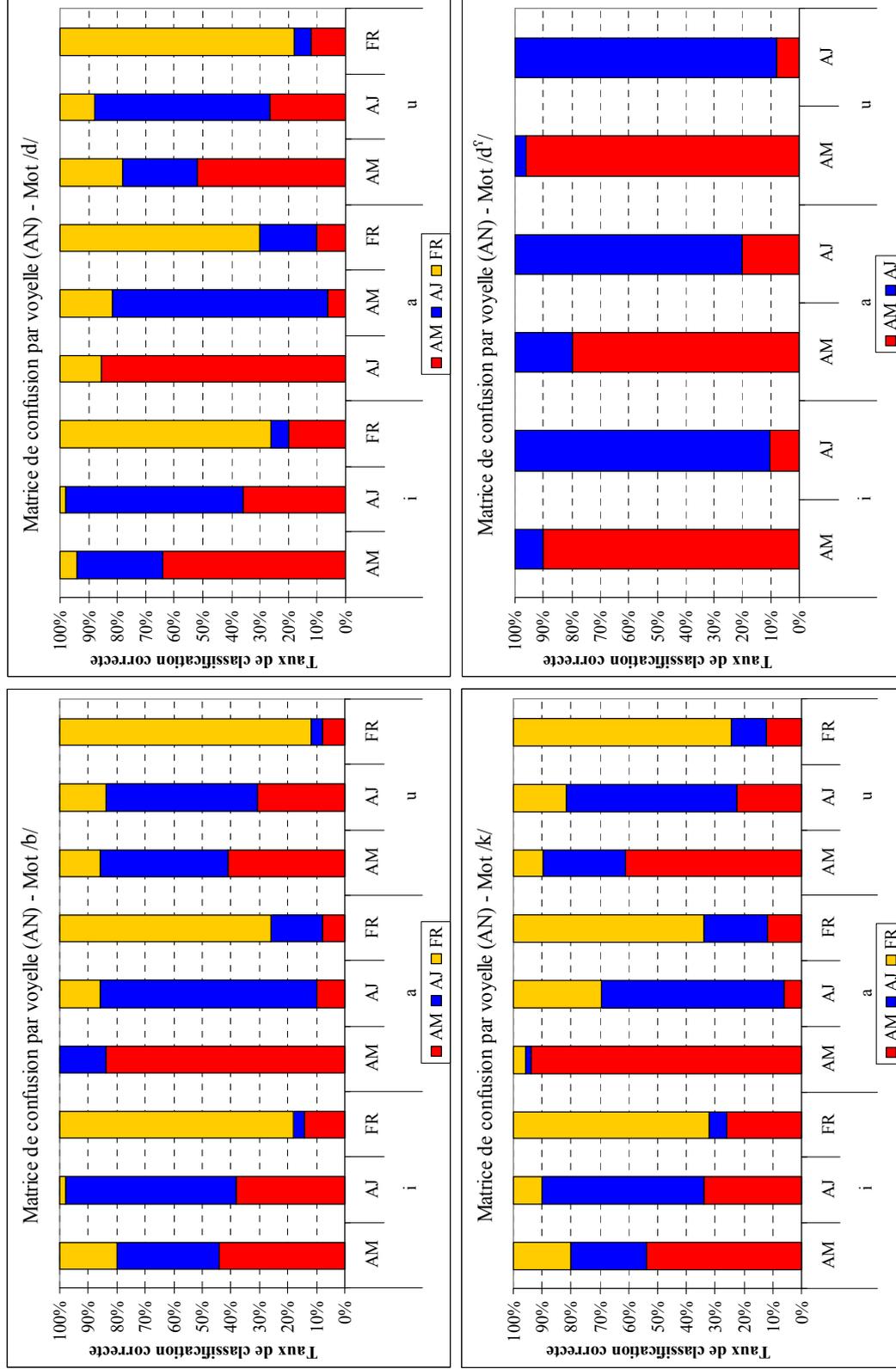


Figure 60 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression linéaire) avec normalisation du temps (AN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f k/ entre AM & AJ).

Ces résultats montrent que les coefficients de la régression linéaire : les **pent**es **formantiques** et les ordonnées à l'origine, avec les valeurs de durée de la transition sont utiles à la discrimination des trois langues. La normalisation du temps n'a pas apportée d'amélioration à la discrimination des trois langues : le taux de classification correcte a baissé significativement de 12% (BinomialQ ; $p < 0,001$) indiquant ainsi que la durée de la transition est un indice important utilisé par l'analyse discriminante pour classer les trois systèmes. Il semblerait alors que la durée de la transition affecte bel et bien la caractérisation **dynamique** des voyelles /i a u/ entre les trois langues, allant ainsi dans le sens des propositions de Lindblom (1963_a). Toutefois, les taux de classification correcte obtenus restent élevés lorsque la durée de la transition formantique est normalisée, indiquant une différence au niveau des points de départ des transitions, des degrés d'inclinaison des **pent**es **formantiques** et des différences de « cibles vocaliques » dérivées entre les trois langues.

Nous nous intéressons dans les parties qui suivent aux résultats de l'analyse discriminante en utilisant les coefficients des deux régressions polynomiales de 2^{ème} et 3^{ème} ordre. Il est à noter que nous n'avons pas pu présenter visuellement les résultats de ces deux analyses à cause de la difficulté d'obtenir une représentation multidimensionnelle de ces paramètres (voir néanmoins Figure 45). Nous évaluerons ainsi l'apport de ces analyses non linéaires dans la discrimination entre les trois langues.

4.3.4.2.2. Régression polynomiale de 2^{ème} ordre

Les paramètres qui ont été utilisés pour cette analyse sont les coefficients de la régression polynomiale de 2^{ème} ordre⁶⁰ de F_1 , F_2 & F_3 des voyelles /i a u/, avec ou sans normalisation du temps. Les **pent**es **formantiques** différentes entre les trois langues ont contribué à la séparation des voyelles dans chacune des langues. Dans le Tableau 16, nous présentons les taux de classification correcte des trois langues en fonction de chaque situation de production, de chaque lieu d'articulation et de chaque voyelle. Les taux de classification correcte montrent la possibilité de discriminer les trois langues en fonction des différents lieux d'articulation et/ou de la situation de production. La normalisation du temps baisse

⁶⁰ Les coefficients sont : a_0 = la moyenne des valeurs, a_1 = la droite de régression et a_2 = l'ordonnée à l'origine.

significativement, le taux moyen de classification des trois langues par voyelle de 10,25% (64,75% avec normalisation et 75,0% sans normalisation), (BinomialQ ; $p < 0,001$).

Tableau 16 : Taux de classification correcte des trois langues à partir de l'analyse de polynomiale de 2^{ème} ordre, avec et sans normalisation du temps, en fonction de la situation de production, du lieu d'articulation (/b d k/ entre les trois langues et /d^ʕ/ entre les deux dialectes arabes) et des voyelles /i a u/. Les valeurs entre parenthèses indiquent les taux avec normalisation du temps (Taux significatifs, Box's M ; $p < 0,001$).

		/b/	/d/	/k/	/d ^ʕ /
Mot	/i/	78,0% (60,7%)	74,7% (70,7%)	64,7% (60,0%)	92,9% (92,9%)
	/a/	88,0% (78,7%)	87,8% (73,6%)	93,2% (74,3%)	85,0% (84,0%)
	/u/	70,3% (62,8%)	69,8% (59,7%)	71,4% (64,6%)	94,0% (96,0%)
Syllabe	/i/	73,3% (65,3%)	74,0% (68,0%)	72,0% (58,7%)	78,6% (76,5%)
	/a/	74,5% (58,4%)	74,3% (59,5%)	79,1% (70,3%)	84,0% (75,0%)
	/u/	75,0% (68,9%)	72,5% (70,5%)	79,7% (71,6%)	79,0% (72,0%)
Isolation	/i/	64,0% (56,7%)	71,3% (55,3%)	77,9% (64,4%)	70,4% (67,3%)
	/a/	75,3% (62,0%)	76,4% (68,2%)	72,3% (58,8%)	78,6% (73,5%)
	/u/	74,3% (64,9%)	73,2% (56,4%)	68,0% (65,3%)	77,8% (73,7%)

Les taux moyens de classification de l'analyse discriminante sont élevés (proche des 76%) avec des taux variables. Par exemple, le taux le plus bas est de 64% lors de la classification de la voyelle /i/ dans l'environnement /b/ en Isolation (sans normalisation du temps) et le plus haut est de 96% pour la voyelle /u/ en contexte /d^ʕ/ en Mot (avec la normalisation du temps). Ces sont les confusions dans les classification des voyelles dans les trois langues qui ont baissé ces taux. Afin de les examiner, nous les présentons sur les Figures 61 et 62 en situation de production Mot (pour les deux autres situations, voir Annexes 10 et 11). Comme pour les deux autres analyses (en **statique** et en **dynamique** avec la régression linéaire), nous observons des confusions plus marquées dans la classification avec la normalisation du temps et ce quels que soient le lieu d'articulation et la situation de production. Le taux moyen de la classification des voyelles /i a u/ en AM et AJ dans les trois situations de production et dans l'environnement /d^ʕ/ est très important, indiquant que ce sont les différences quantifiées directement par les **pentés formantiques** qui sont à l'origine de cette différence, en **statique** en revanche, les confusions étaient plus importantes.

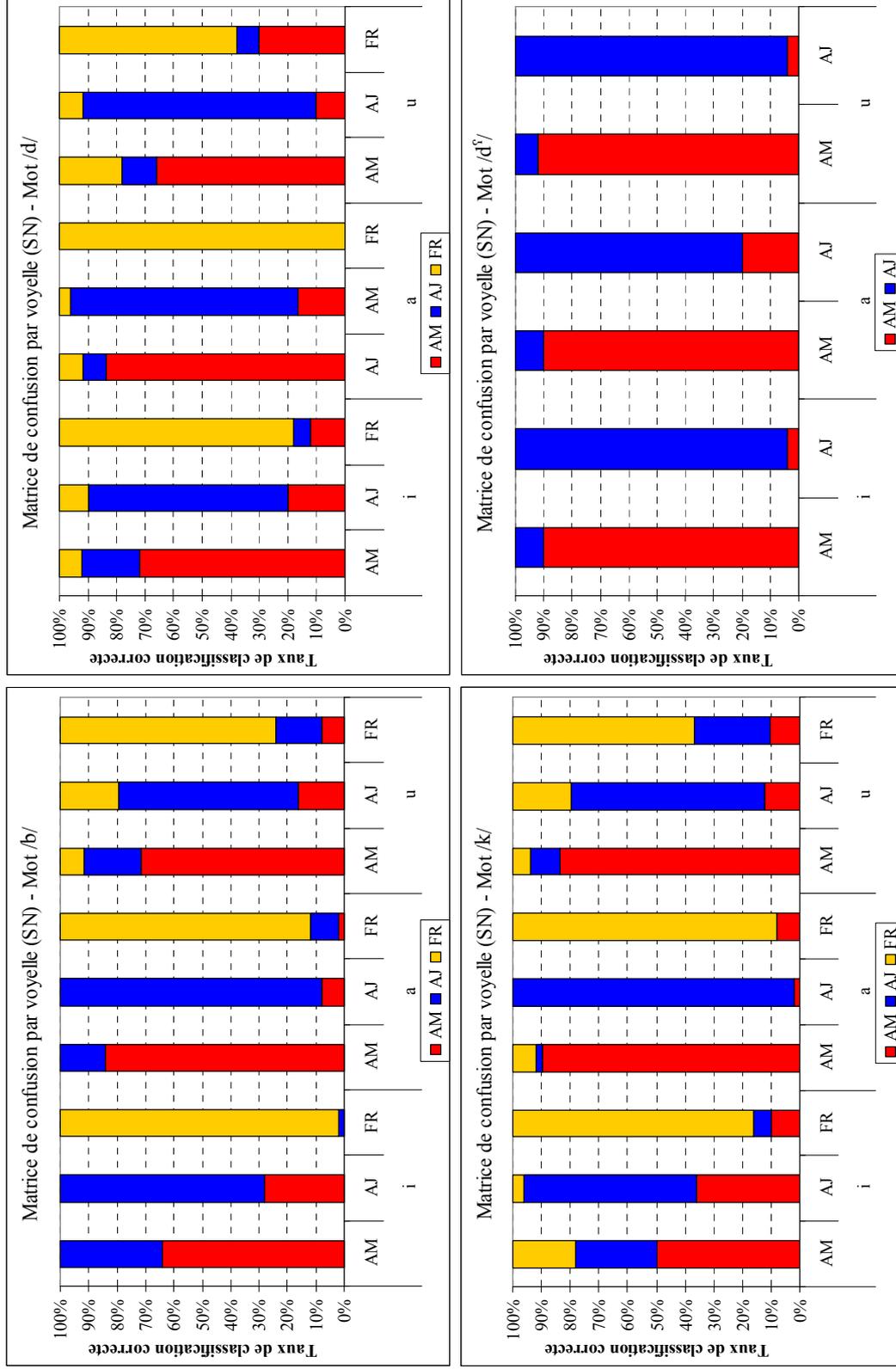


Figure 61 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression polynomiale 2^{ème} ordre) sans normalisation du temps (SN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f entre AM & AJ).

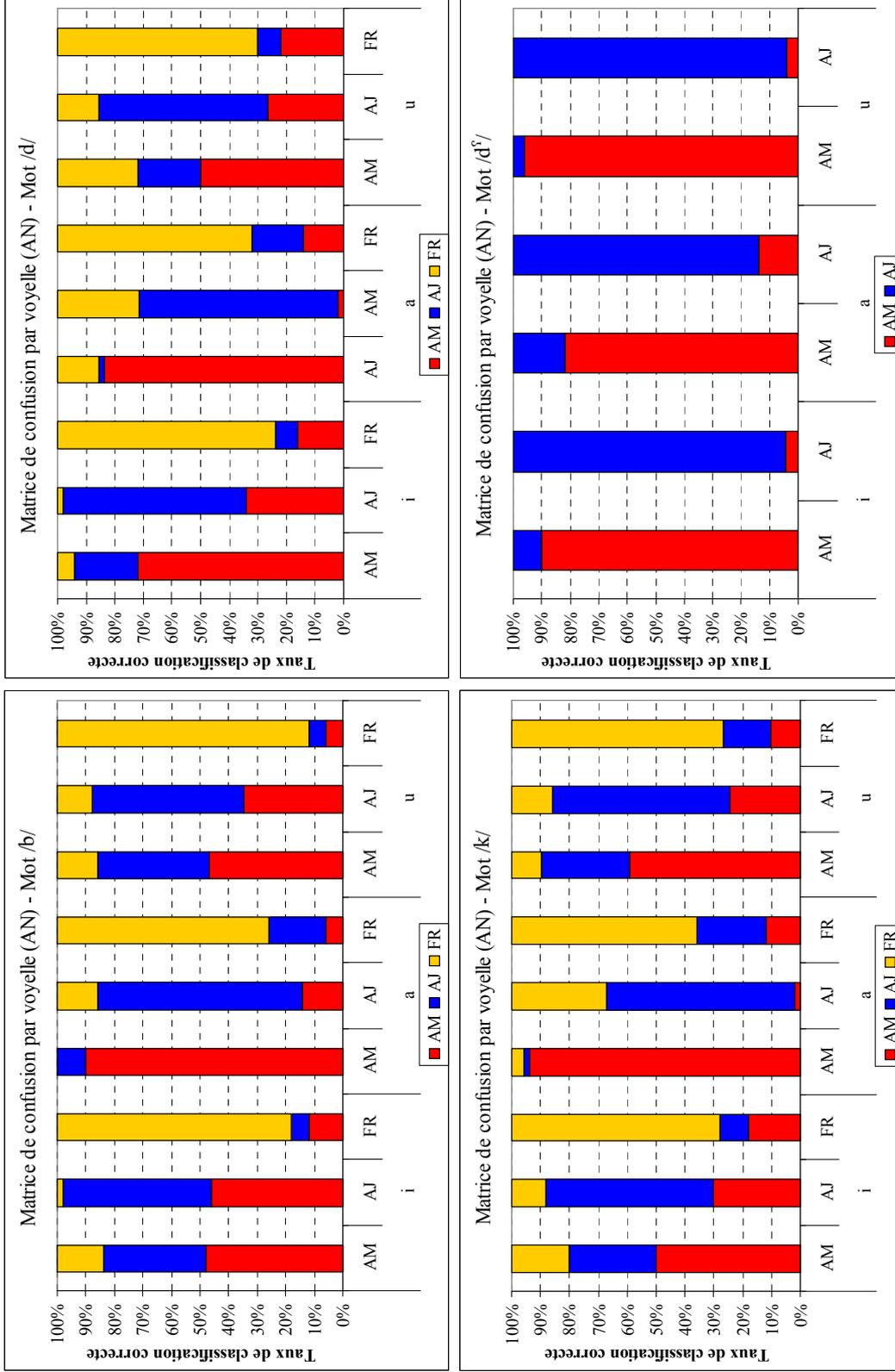


Figure 62 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression polynomiale 2^{ème} ordre) avec normalisation du temps (AN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f/ entre AM & AJ).

4.3.4.2.3. Régression polynomiale de 3^{ème} ordre

Les paramètres des coefficients de la régression polynomiale de 3^{ème} ordre⁶¹ de F_1 , F_2 & F_3 des voyelles /i a u/, ainsi que la durée de la transition formantique lorsque le temps n'est pas normalisé. Les taux de classification par voyelle sont présentés dans le Tableau 17

Tableau 17 : Taux de classification correcte des trois langues à partir de l'analyse de régression polynomiale de 3^{ème} ordre, avec et sans normalisation du temps, en fonction de la situation de production, du lieu d'articulation (/b d k/ entre les trois langues et /d^ʕ/ entre les deux dialectes arabes) et des voyelles /i a u/. Les valeurs entre parenthèses indiquent les taux avec normalisation du temps (Taux significatifs, Box's M ; $p < 0,001$).

		/b/	/d/	/k/	/d ^ʕ /
Mot	/i/	77,3% (74,7%)	74,0% (68,0%)	63,3% (62,7%)	91,8% (90,8%)
	/a/	88,0% (70,0%)	87,8% (75,7%)	89,2% (75,7%)	81,0% (81,0%)
	/u/	73,0% (64,2%)	71,8% (65,1%)	72,1% (61,2%)	93,0% (94,0%)
Syllabe	/i/	75,3% (90,7%)	75,3% (73,3%)	72,0% (57,3%)	80,6% (74,5%)
	/a/	75,2% (53,7%)	73,0% (55,4%)	79,1% (72,3%)	83,0% (77,0%)
	/u/	75,0% (66,2%)	69,8% (67,1%)	79,7% (67,6%)	77,0% (75,0%)
Isolation	/i/	66,0% (84,0%)	69,3% (66,0%)	77,9% (63,8%)	67,3% (64,3%)
	/a/	74,0% (66,0%)	71,6% (62,2%)	71,6% (61,5%)	77,6% (73,5%)
	/u/	73,0% (64,9%)	68,5% (57,0%)	68,7% (67,3%)	75,8% (75,8%)

Le taux moyen de classification est proche des 75%. Lorsque la durée de la transition formantique est normalisée, nous obtenons une baisse significative du taux moyen de la classification des trois langues de 7,33% (67,17% avec normalisation et 74,5% sans normalisation), (BinomialQ ; $p < 0,001$). Les taux moyens de classification de l'analyse discriminante sont relativement élevés, mais les confusions dans la classification sont plus importantes par rapport à l'analyse de régression polynomiale de 2^{ème} ordre. En effet, ces confusions concernent les trois langues et non seulement l'AM et l'AJ et plus particulièrement lorsque le temps est normalisé. Les résultats graphiques des confusions de classification en situation Mot sont présentés sur les Figures 63 et 64 (voir Annexes 12 et 13 pour les graphiques en situations Syllabe et Isolation).

⁶¹ Les coefficients sont : a_0 = la moyenne des valeurs, a_1 = la droite de régression, a_2 = l'excursion parabolique de la voyelle indépendamment de sa droite de régression et a_3 = l'ordonnée à l'origine.

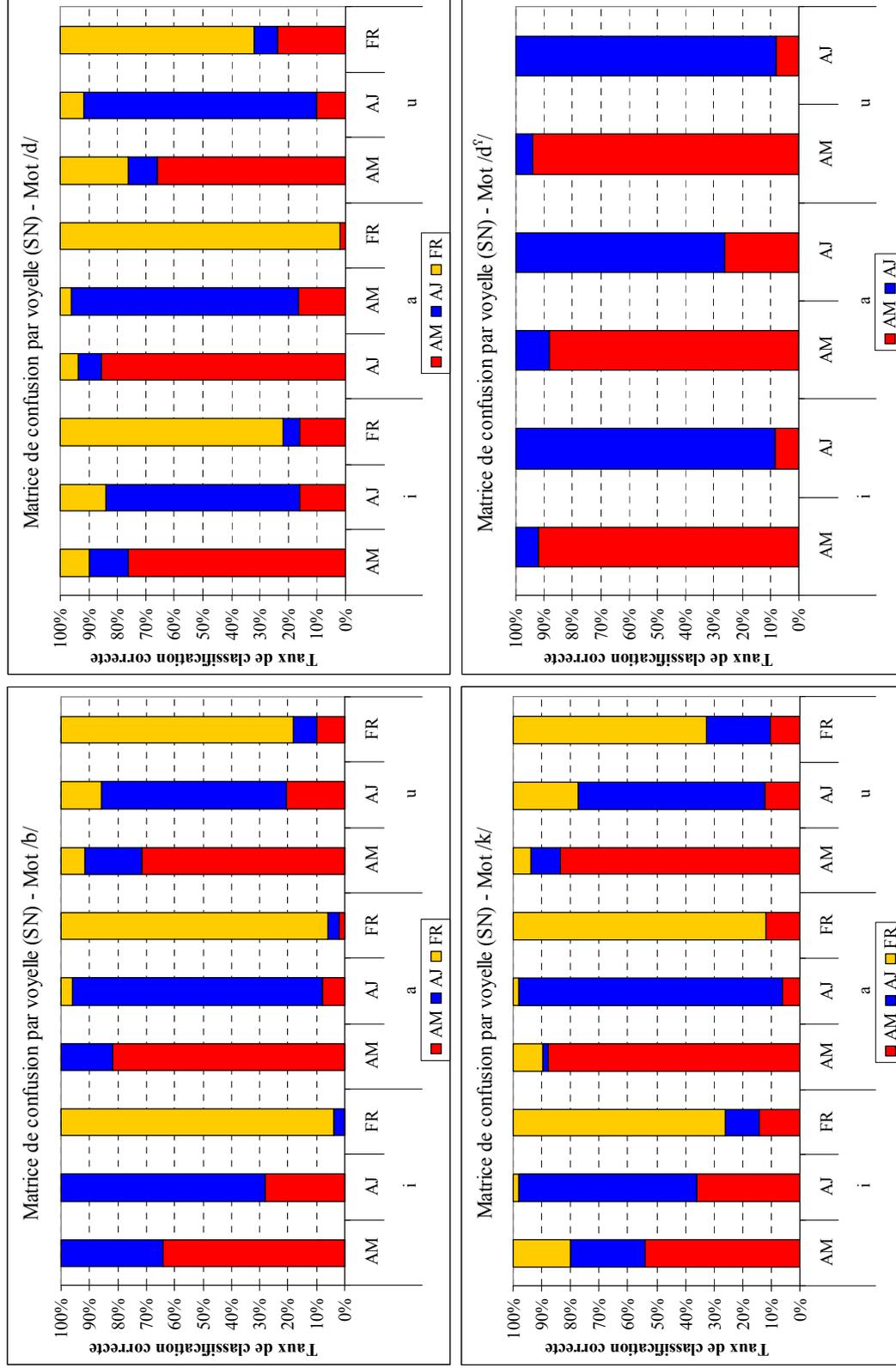


Figure 63 : Matrices de confusion de classification de chaque voyelle /i a u/ en dynamique (régression polynomiale 3^{ème} ordre) sans normalisation du temps (SN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f/ entre AM & AJ).

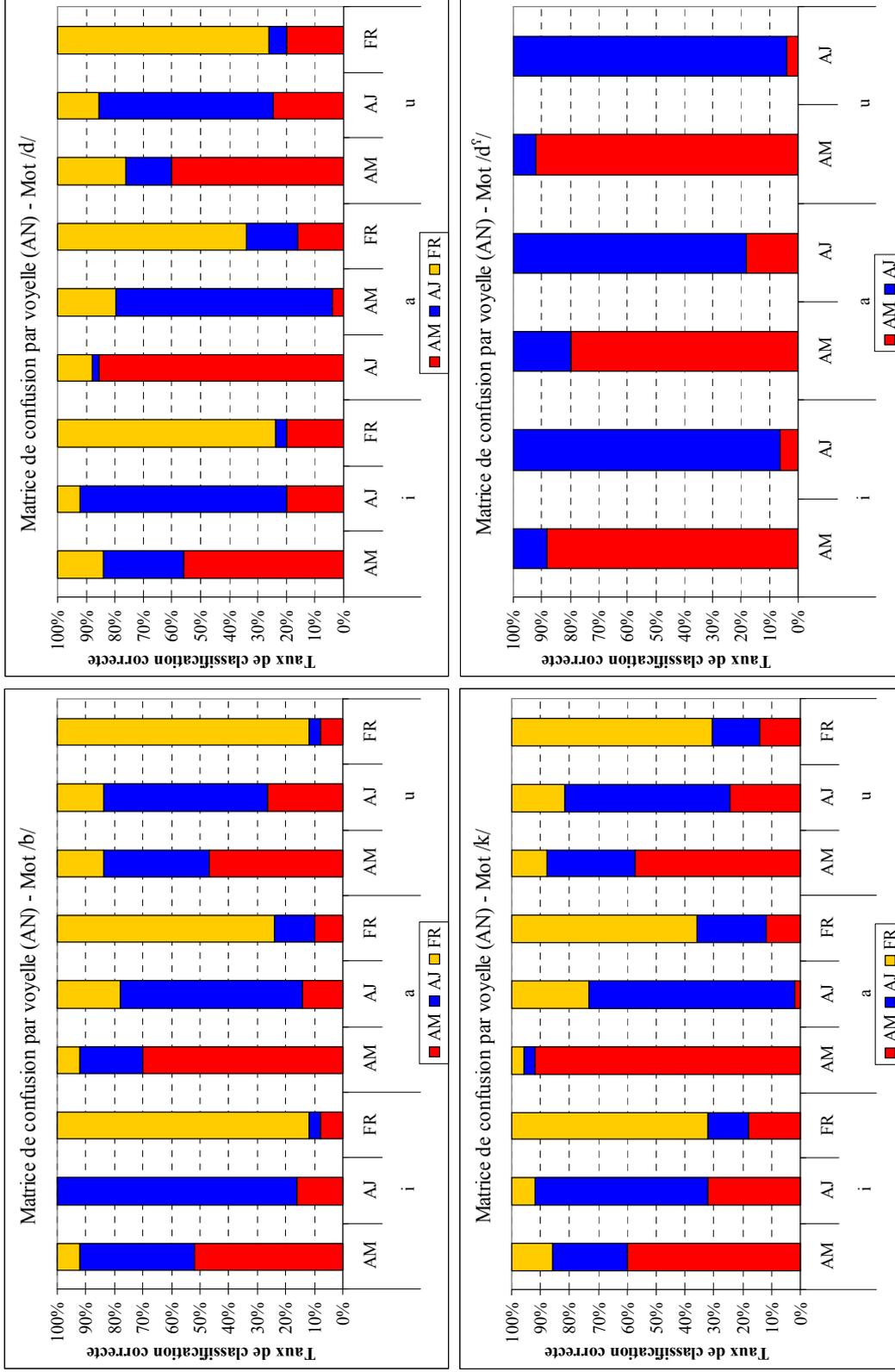


Figure 64 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression polynomiale 3^{ème} ordre) avec normalisation du temps (AN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f/ entre AM & AJ).

Lorsque le temps n'est pas normalisé, les voyelles en FR sont relativement bien classifiées avec des taux moyens proches des 70% ; les confusions de classification quant à elles sont plus importantes qu'avec les autres analyses discriminantes. Les voyelles du FR ont été souvent confondues avec celles de l'AJ et très peu (voire pas du tout) avec les voyelles de l'AM. Une baisse des taux de classification est observée et plus particulièrement lorsque le temps est normalisé. En effet, les confusions de la classification sont plus importantes que lorsque le temps n'est pas normalisé. En contexte /d^ɛ/ nous observons une baisse du taux de classification comparée à l'analyse discriminante avec les coefficients de la régression polynomiale de 2^{ème} ordre. Dans les deux autres situations de production, nous obtenons également plus de confusions, avec et sans normalisation du temps.

Nous venons d'examiner les taux de classification correcte de l'analyse discriminante obtenus en **statique** (à partir des « cibles vocaliques ») et en **dynamique** (à partir des **pentés formantiques** obtenues avec les coefficients des régressions linéaire et polynomiales de 2^{ème} et de 3^{ème} ordre). Ces taux relativement élevés montrent qu'il est possible de discriminer les voyelles des trois langues à partir des différents paramètres proposés, confirmant ainsi les résultats obtenus sur les effets de la densité des systèmes.

Dans cette partie, nous nous sommes intéressés à la comparaison inter-langues à partir des trois voyelles communes /i a u/. Nous avons montré que les **pentés formantiques** obtenues à partir des régressions linéaires et polynomiales (de 2^{ème} et 3^{ème} ordre) permettent de quantifier les différences observées en fonction des points de départ des transitions formantiques ; des « cibles vocaliques » dérivées et des changements formantiques observés dans les trois langues en fonction de la situation de production et des lieux d'articulation des consonnes adjacentes. Ces différences ont été validées avec l'aide d'une méthode de classification des voyelles à partir d'une analyse discriminante. La prochaine étape consistera à la quantification des différences observées en fonction de chaque langue. Nous présenterons ainsi dans les parties qui suivent les spécificités de chaque langue en **statique** et en **dynamique** pour chaque voyelle, et plus particulièrement pour les voyelles problématiques

dans chaque langue, à savoir les voyelles brèves en AM et AJ et les voyelles dites « médianes » et /a a/ en FR.

4.4. SPÉCIFICITÉS DE CHAQUE LANGUE

4.4.1. *En Arabe Marocain*

Nous nous intéressons dans cette partie aux résultats obtenus en AM en **statique** et en **dynamique**. Nous commencerons par décrire le système vocalique de l'AM à partir des « cibles vocaliques » obtenues dans les trois situations de production et pour les quatre consonnes /b d d^f k/. Nous évaluerons le statut des voyelles brèves en AM, à savoir une seule voyelle centrale /ə/ ou deux (/ə u/), (pour plus de détails, voir § 1.4.2). Une fois les résultats en **statique** développés, nous comparerons les résultats des **pentés formantiques** des voyelles de l'AM, et plus particulièrement en fonction des voyelles brèves. Nous évaluerons enfin les résultats de l'analyse discriminante qui va nous permettre de valider ces résultats.

4.4.1.1. Résultats en statique

L'espace vocalique moyen en AM est affecté par la situation de production (voir Figure 65) sur F_1 ($F(2, 2892)=23,02$; $p<0,001$) et sur F_2 ($F(2, 2892)=92,37$; $p<0,001$) : il est plus « périphérique » en situation Isolation et plus « centralisé » en Mot ; en Syllabe, l'espace vocalique est intermédiaire (moyenne d'aire de dispersion en Mot : 6,15 Bark² ; en Syllabe : 7,62 Bark² et en Isolation : 8,59 Bark², ($F(2, 108)=14,62$; $p<0,001$)).

Les voyelles produites en AM sont significativement différentes sur les deux axes, i.e., le système vocalique attesté est composé des cinq voyelles /i: a: ə u u:/ (sur F_1 ($F(4, 2892)=1766,43$; $p<0,001$) et sur F_2 ($F(2, 2892)=7210,08$; $p<0,001$)). Plus spécifiquement, les timbres vocaliques obtenus pour les deux voyelles brèves /ə u/ de l'AM sont significativement différents. Bien qu'elles soient « centralisées » toutes les deux, /u/ est significativement plus « fermée » sur F_1 ($F(1, 1140)=15,31$; $p<0,001$) et plus « postériorisée » sur F_2 ($F(1, 1140)=508,13$; $p<0,001$) que la voyelle /ə/. L'interaction entre la situation de production et les voyelles est significative : sur F_1 , les voyelles /i: u:/ produites

en Isolation sont significativement plus « fermées » et les voyelles /a: ə u/ sont plus « ouvertes » par rapport à celles produites en Mot ($F(8, 2892)=26,23$; $p<0,001$). Sur F_2 , la voyelle /i:/ produite en Isolation est significativement plus antérieure que celle en Mot ; /a: u:/ sont significativement plus postérieures ($F(8, 2892)=92,07$; $p<0,001$) ; aucune différence n'est observée pour les voyelles /ə u/ ni sur F_1 ($F(1, 1140)=1,37$; $p=0,25$), ni sur F_2 ($F(1, 1140)=0,34$; $p=0,71$).

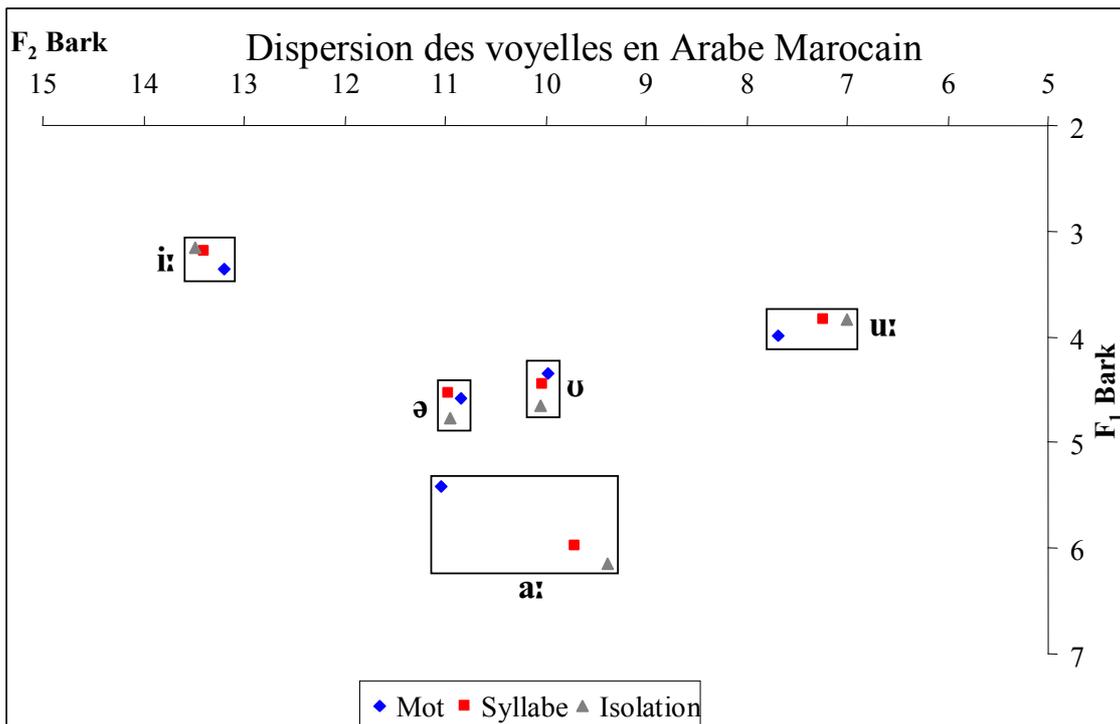


Figure 65 : Espace vocalique en AM en fonction de la situation de production⁶².

Les espaces vocaliques obtenus pour chaque situation de production et chaque consonne /b d d^s k/ sont présentés sur la Figure 66. Les résultats statistiques montrent globalement que l'espace vocalique obtenu en Isolation est significativement plus « grand » et significativement plus « petit » en Mot (sur F_1 : $F(6, 2892)=20,67$; $p<0,001$; et sur F_2 : $F(6, 2892)=51,43$; $p<0,001$).

⁶² Les carrés dessinés autour de chaque catégorie vocalique sont affichés simplement par souci de visibilité pour regrouper les données par timbre vocalique. Leur taille est arbitraire et donc sans lien avec la dispersion des ellipses.

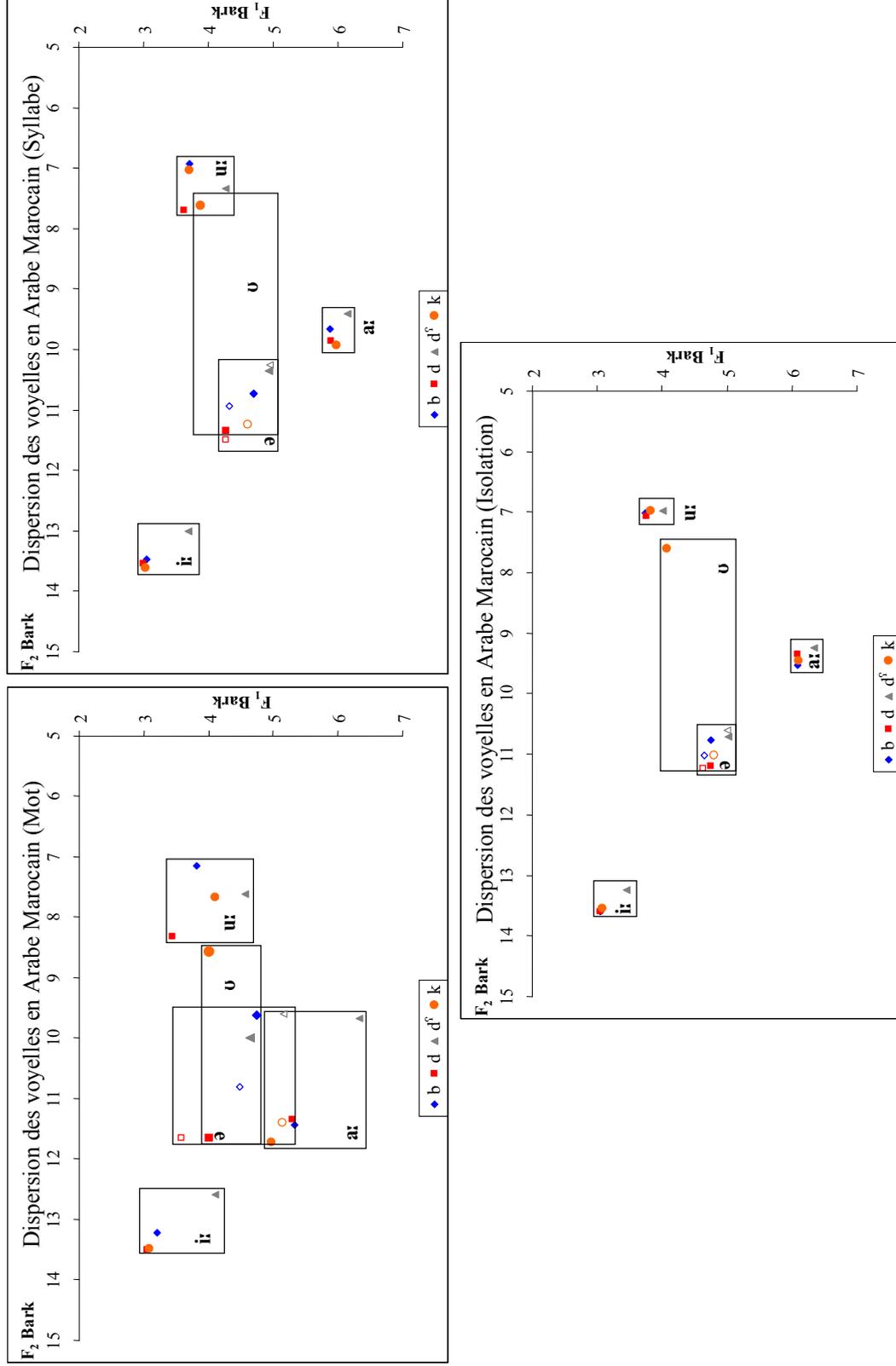


Figure 66 : Espaces vocaliques en AM obtenus en fonction de chaque situation de production et du lieu d'articulation des consonnes /b d d' k/. Les symboles utilisés pour la voyelle /ə/ sont ceux avec un fond vide.

Nous obtenons également des différences liées aux timbres des voyelles produits par les locuteurs marocains en fonction de la situation de production et du lieu d'articulation et plus particulièrement pour les deux voyelles brèves. Sur F_1 , et en situations Mot et Syllabe, la voyelle /ə/ est significativement plus « fermée » dans les environnement /b d/ ; dans les environnements /d^h k/, c'est /u/ qui est plus « fermée » ; en Isolation, /ə/ est significativement plus « fermée » dans les environnement /b d d^h/ ; en revanche, dans l'environnement /k/, /u/ est plus « fermée », ($F(6, 1140)=3,37$; $p<0,001$). Sur F_2 , et en situation Mot, la voyelle /ə/ est significativement plus « antérieure » dans les environnement /b k/ ; dans les environnements /d d^h/, /u/ est plus « antérieure » ; en Syllabe et Isolation, /ə/ est significativement plus « antérieure » dans les environnement /b d k/ ; dans l'environnement /d^h/, /u/ est plus « antérieure », ($F(6, 1140)=8,28$; $p<0,001$). Les différences obtenues montrent l'existence de deux timbres vocaliques brefs distincts en AM.

En ce qui concerne les voyelles longues, les résultats statistiques montrent que c'est dans l'environnement /d^h/ que ces voyelles sont significativement plus « ouvertes » et plus « postérieures » conformément aux données de la littérature sur les effets de la pharyngalisation sur la dispersion des voyelles (voir Ghazeli, 1977, 1981 ; Elgendy, 2001 ; Al-Tamimi & Barkat-Defradas, 2002 ; etc...).

Bien que les aires de dispersion des voyelles produites en AM soient significativement différentes en fonction de la situation de production, l'interaction entre la situation de production et le lieu d'articulation n'est pas significative, indiquant que la taille plus grande de l'aire en situation Isolation par rapport à celle en Mot est identique quel que soit le lieu d'articulation ($F(6, 108)=0,80$; $p=0,57$), (voir Figure 67).

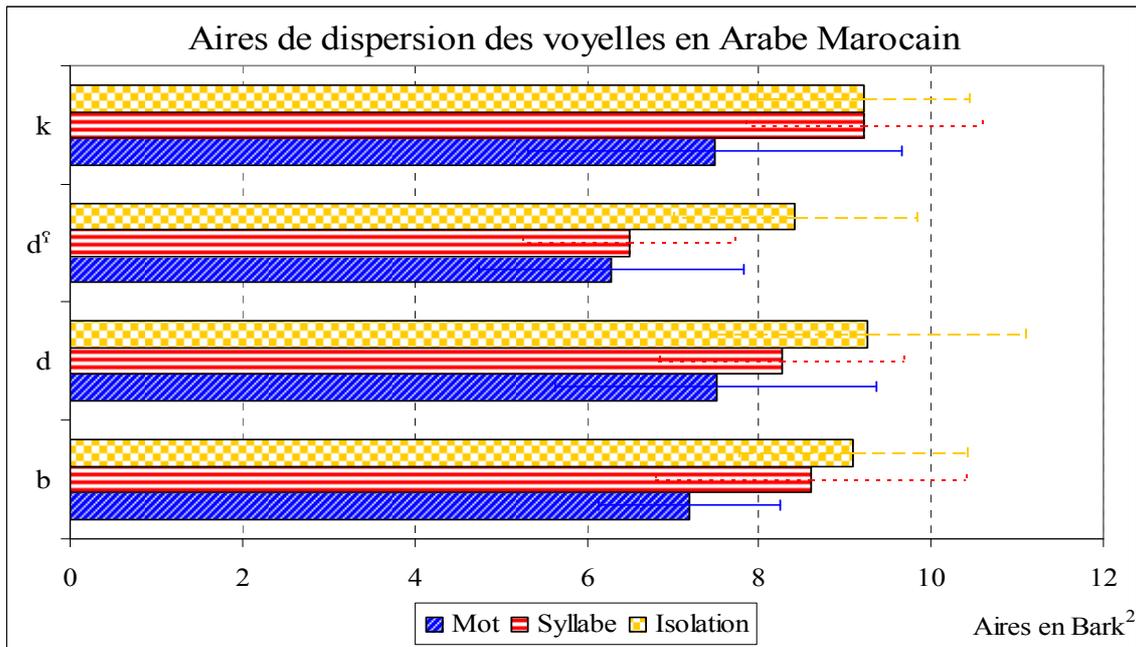


Figure 67 : Aires de dispersion (moyenne et écart-type) des voyelles de l'AM en fonction de la situation de production et du lieu d'articulation.

La taille des ellipses de chaque voyelle en fonction de la situation de production et du lieu d'articulation a été calculée (voir Figure 68 et Tableau 18). Les résultats montrent des ellipses de dispersion plus importantes pour les voyelles /i: a: u:/ en Mot, faibles en Syllabe et intermédiaires en Isolation ; les deux voyelles /ə ʊ/ présentent des ellipses plus grandes que celles obtenues pour les voyelles /i: a: u:/. En Mot, la voyelle /ʊ/ obtient l'ellipse la plus grande par rapport à toutes les autres voyelles, indiquant que les productions des locuteurs marocains sont significativement différentes.

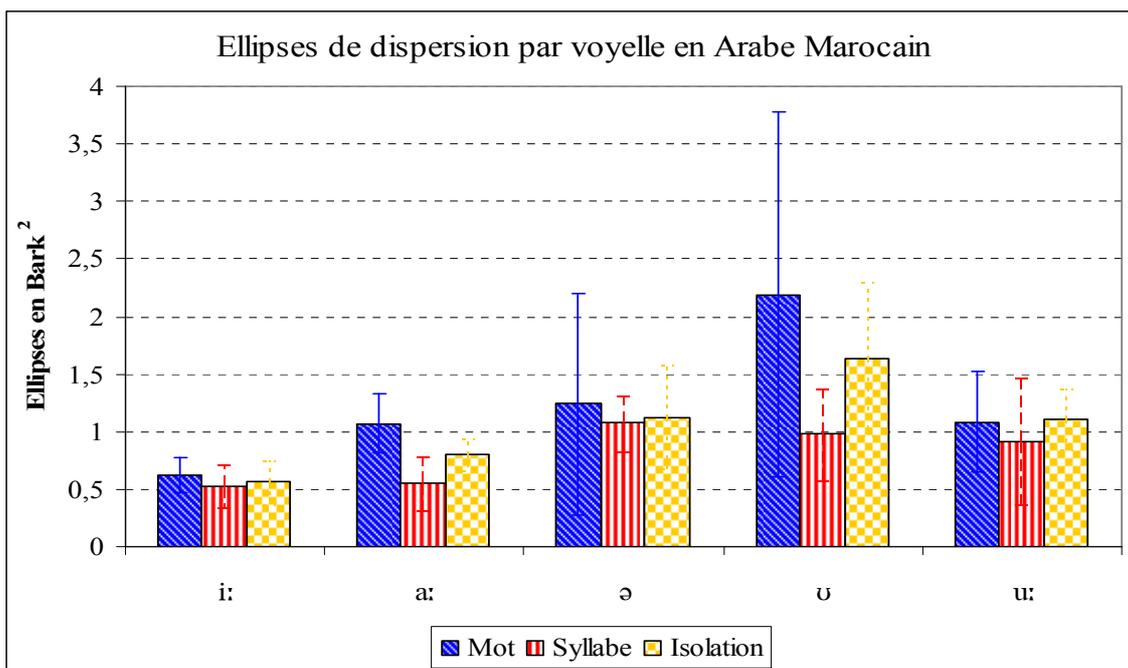


Figure 68 : Ellipses vocaliques en AM en fonction de la situation de production.

Tableau 18 : Taille des ellipses vocaliques en fonction de la situation de production et du lieu d'articulation en AM (en Bark²).

(Bark²)		i:	a:	ə	u	u:
Mot	b	0,65	1,39	0,42	4,23	0,94
	d	0,58	1,10	1,16	0,62	0,59
	d^ʕ	0,82	0,76	0,77	2,60	1,62
	k	0,45	1,03	2,61	1,31	1,20
Syllabe	b	0,46	0,52	1,01	1,03	0,52
	d	0,50	0,90	0,77	0,64	0,75
	d^ʕ	0,80	0,47	1,32	1,52	1,72
	k	0,36	0,34	1,20	0,73	0,67
Isolation	b	0,51	0,82	0,66	0,76	1,16
	d	0,44	0,92	1,27	2,07	0,75
	d^ʕ	0,85	0,87	1,69	2,23	1,37
	k	0,45	0,60	0,87	1,49	1,16

En observant les résultats des tailles des ellipses de dispersion des voyelles /ə/ et /u/, nous obtenons des ellipses très grandes dans tous les contextes consonantiques et quel que soit la situation de production. Ceci montre l'existence d'une variabilité inter-individuelle importante lors de la réalisation de ces voyelles. Nous présentons sur les Figures 69 à 71, la dispersion des voyelles en AM à partir des données individuelles.

Nous observons des variations inter-individuelles importantes au niveau de la réalisation des trois voyelles /i: a: u:/ dans les trois situations de production. Lorsque nous nous intéressons aux deux voyelles brèves de l'AM ; /ə ʊ/, nous observons une séparation complète entre ces deux voyelles en contexte Mot dans les environnements /b k/ : les locuteurs marocains produisent deux voyelles distinctes ; dans les deux autres environnements /d d^h/, ces deux voyelles fusionnent et sont ainsi produites comme un schwa (voir Figure 69). Dans les situations de production Syllabe (Figure 70) et Isolation (Figure 71), les deux voyelles /ə ʊ/ fusionnent dans les environnements /b d d^h/ et sont réalisées comme un schwa, tandis que dans l'environnement vélaire /k/, nous obtenons une réalisation en deux voyelles distinctes. Il semblerait également que ces deux voyelles constituent deux catégories distinctes chez certains locuteurs marocains et une seule et unique pour d'autres.

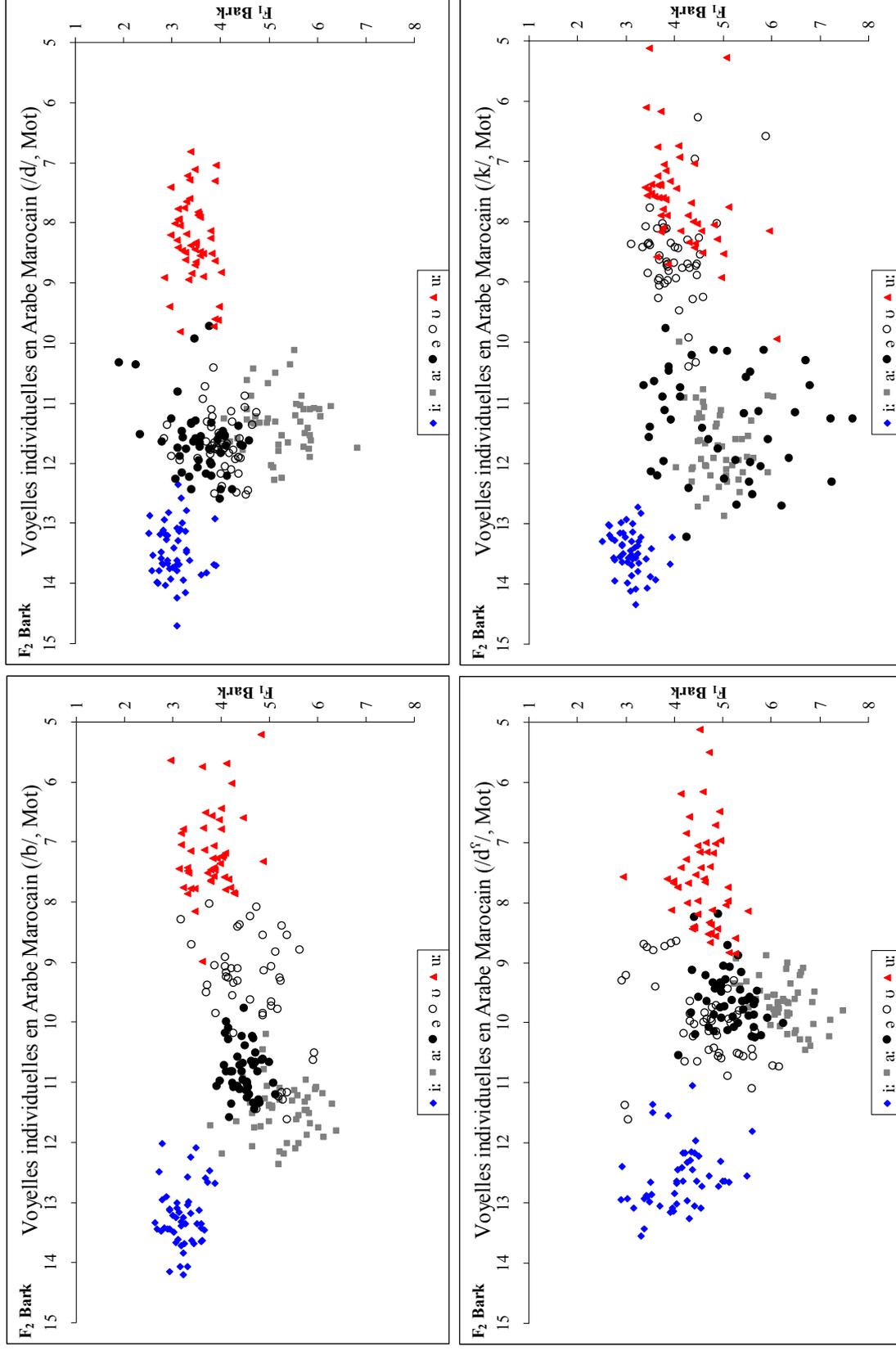


Figure 69 : Dispersion des voyelles en AM à partir des données individuelles, par lieu d'articulation, en Mot.

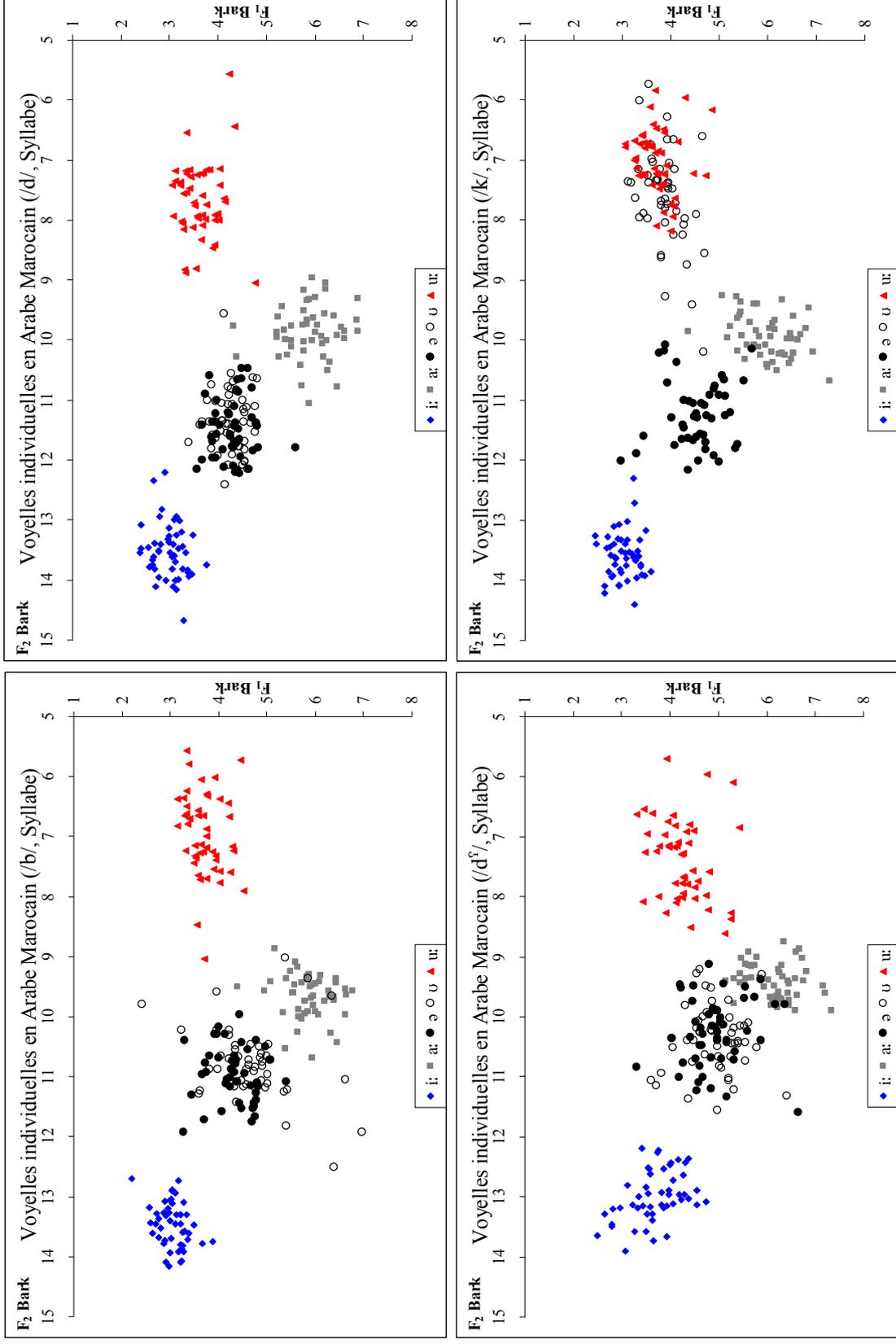


Figure 70 : Dispersion des voyelles en AM à partir des données individuelles, par lieu d'articulation, en Syllabe.

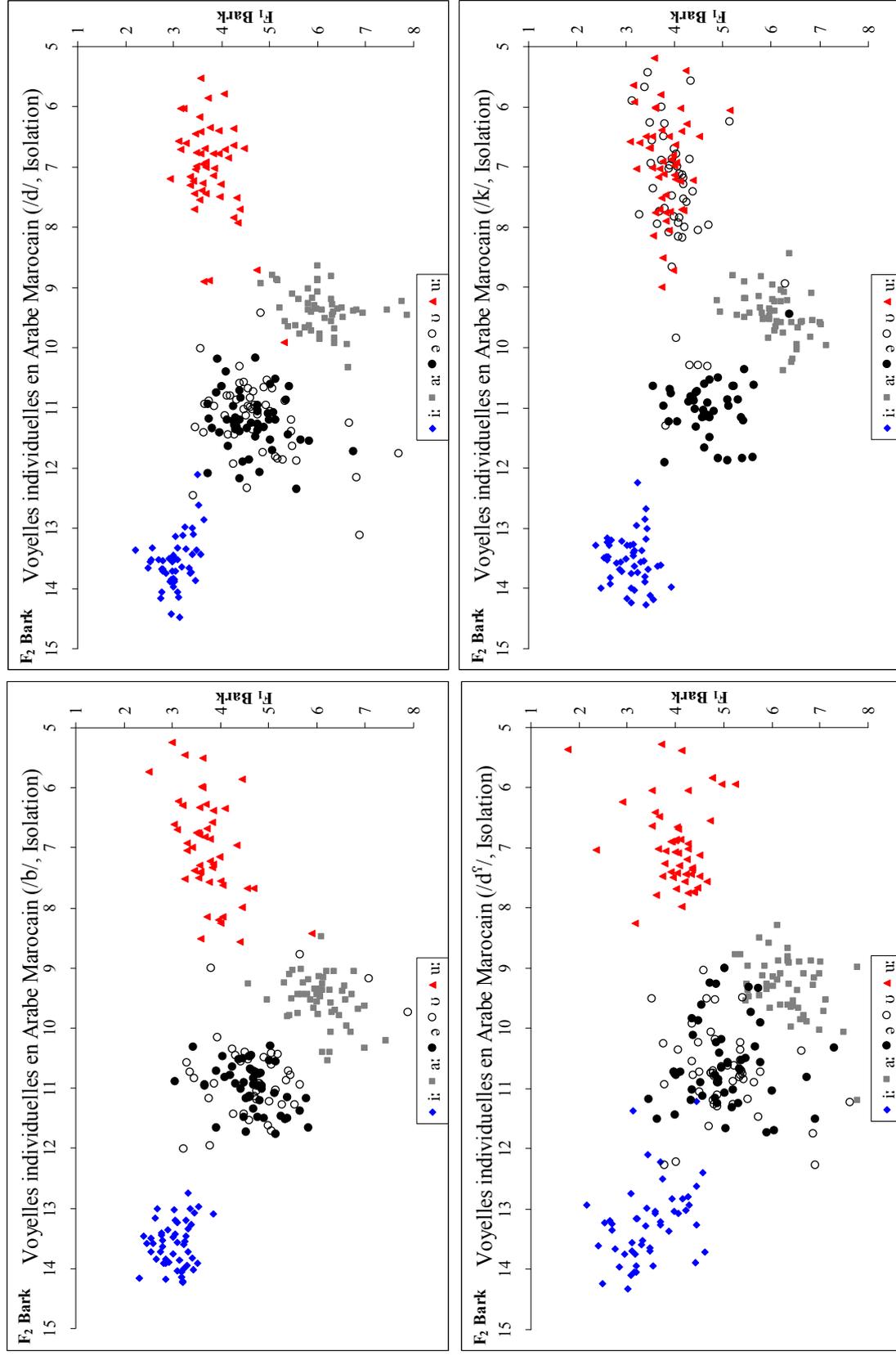


Figure 71 : Dispersion des voyelles en AM à partir des données individuelles, par lieu d'articulation, en Isolation.

4.4.1.2. Résultats en dynamique

Les **pentés formantiques** de F_1 , F_2 & F_3 obtenues en AM sont significativement différentes selon la situation de production : elles sont significativement plus « aplaties » en Isolation et plus « raides » en Mot et Syllabe ($p < 0,001$). Nous présentons sur les Figures 72 à 74 les **pentés formantiques** moyennes en AM en fonction des trois premiers formants de chaque voyelle en fonction du lieu d'articulation, dans la situation de production Mot (pour les résultats graphiques des **pentés formantiques** en Syllabe et Isolation, voir Annexe 14).

En ce qui concerne les **pentés formantiques** obtenues sur F_1 en Mot (voir Figure 72), les résultats montrent une variation en fonction du point de départ de chaque voyelle : toutes les voyelles en AM commencent dans une zone située entre 3 et 4 Bark dans les environnements /b d k/ ; dans l'environnement pharyngalisé /d^ɣ/, le début de transition est plus haut autour des 4 à 5 Bark. Nous savons que les voyelles produites dans l'environnement pharyngalisé /d^ɣ/ sont plus « ouvertes » que dans les autres environnements non pharyngalisés ; nos résultats montrent que cet effet est bel et bien visible dès le début des transitions formantiques ($F(3, 922) = 7,93$; $p < 0,001$). Le degré d'inclinaison des **pentés formantiques** est fonction de chaque voyelle et chaque lieu d'articulation : en moyenne, /i:/ obtient la **pente** la plus « aplatie » et /ə/ la plus « raide » ($F(4, 922) = 30,27$; $p < 0,001$) ; en ce qui concerne le lieu d'articulation, en moyenne, /b/ obtient la **pente** la plus « raide » et /d^ɣ/ la plus « aplatie » ($F(3, 922) = 7,93$; $p < 0,001$). En ce qui concerne les deux voyelles problématiques en AM : /ə u/, ces deux voyelles n'obtiennent pas le même point de départ de leurs transitions formantiques dans les trois environnements /b d d^ɣ/ ; dans l'environnement /k/, la valeur de l'ordonnée à l'origine est identique ($F(3, 338) = 5,70$; $p < 0,001$). Le degré d'inclinaison des **pentés formantiques** de ces deux voyelles est significativement identique dans les trois environnements /b d d^ɣ/ ($F(3, 338) = 0,90$; $p = 0,44$) ; il est identique entre les deux voyelles /u u:/ dans l'environnement /k/.

Pour F_2 (voir Figure 73), les **pentés formantiques** obtenus sont statistiquement différentes en fonction des voyelles et du lieu d'articulation : en moyenne et en valeur absolue, la voyelle /ə/ obtient la **pente** la plus « raide » tandis que c'est /u/ qui obtient la

pente la plus « aplatie » ($F(4, 922) = 43,76$; $p < 0,001$) ; la **pen**te la plus « raide » est obtenue pour la consonne /d/ et la plus « aplatie » /k/ ($F(3, 922) = 9,06$; $p < 0,001$). En ce qui concerne les deux voyelles /ə u/, des différences importantes sont observées : dans les environnements /b k/, les points de départ des transitions de ces deux voyelles sont significativement différentes ($F(3, 338) = 123,06$; $p < 0,001$), des cibles vocaliques dérivées et des degrés d'inclinaison des **pen**tes **formantiques** ($F(3, 338) = 4,20$; $p < 0,01$) ; tandis que dans les environnements /d d^h/, le même point de départ, la même « cible vocalique » et le même degré d'inclinaison des **pen**tes **formantiques** sont obtenus. Ces résultats vont dans le sens de ceux obtenus en **statique** (voir § 4.4.1) et indiquent la présence de deux voyelles différentes avec des effets coarticulaires observés plus particulièrement sur les timbres des voyelles en contexte /d d^h/.

Quant aux résultats sur F₃ (voir Figure 74), les **pen**tes **formantiques** obtenues sont significativement différentes en fonction des voyelles et des lieux d'articulation : en moyenne, la voyelle /i:/ obtient la **pen**te la plus « raide » tandis que c'est /u:/ qui obtient la **pen**te la plus « aplatie » ($F(4, 922) = 5,84$; $p < 0,001$) ; la **pen**te la plus « raide » est obtenue pour la consonne /d^h/ et la plus « aplatie » /d/ ($F(3, 922) = 7,14$; $p < 0,001$). Les différences entre les **pen**tes **formantiques** obtenues pour les deux voyelles problématiques /ə u/ sont significativement différentes : les degrés d'inclinaison des **pen**tes pour ces deux voyelles sont différents, /u/ obtient la **pen**te la plus « aplatie » ($F(1, 338) = 4,45$; $p < 0,05$), résultat homogène quel que soit le lieu d'articulation ($F(3, 338) = 1,34$; $p = 0,26$) ; les points de départ des ces deux voyelles par lieu d'articulation sont différents ($F(3, 338) = 15,08$; $p < 0,001$).

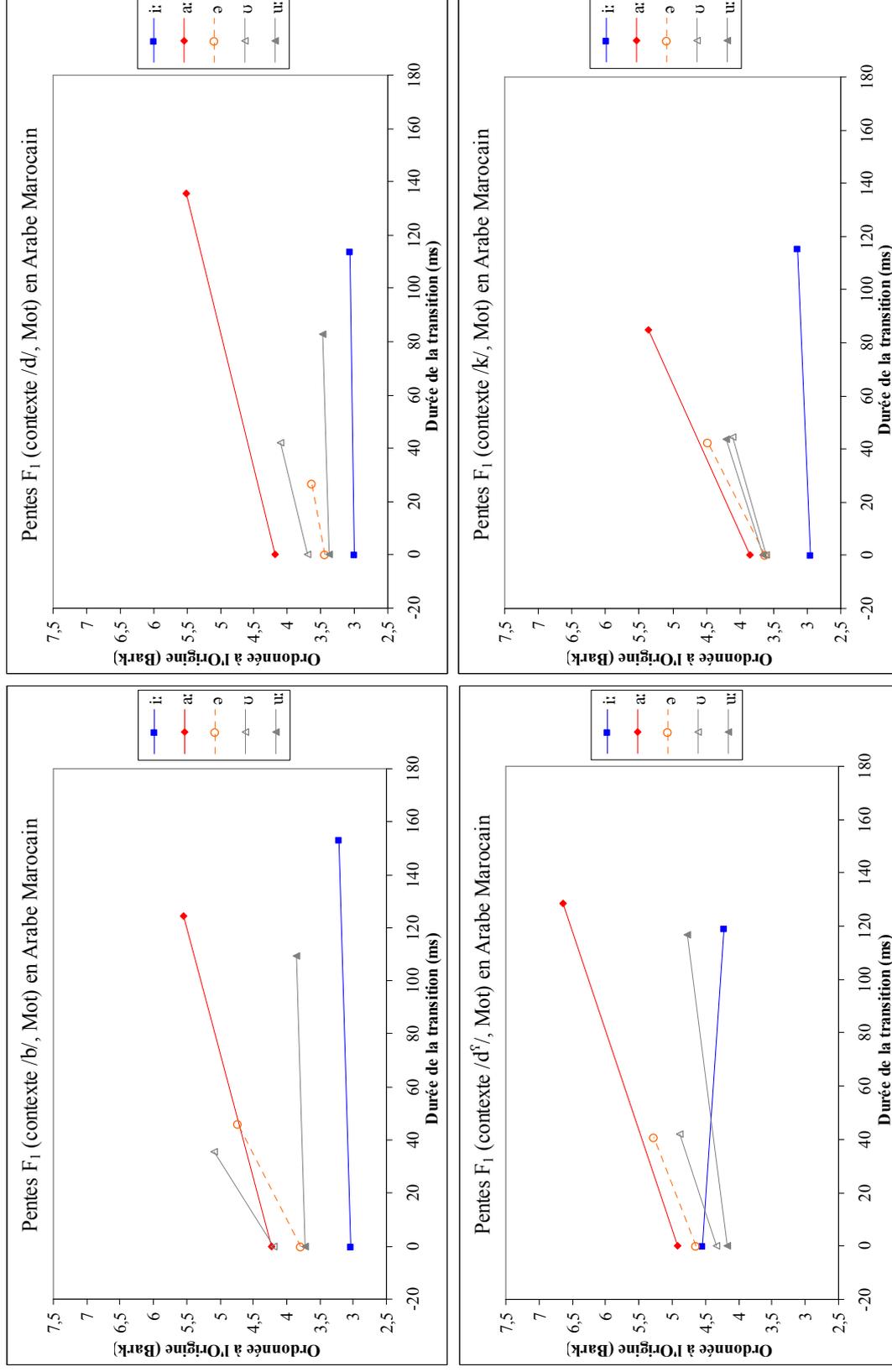


Figure 72 : Pentes formantiques de F₁ des voyelles en AM sans normalisation du temps, en fonction du lieu d'articulation en Mot.

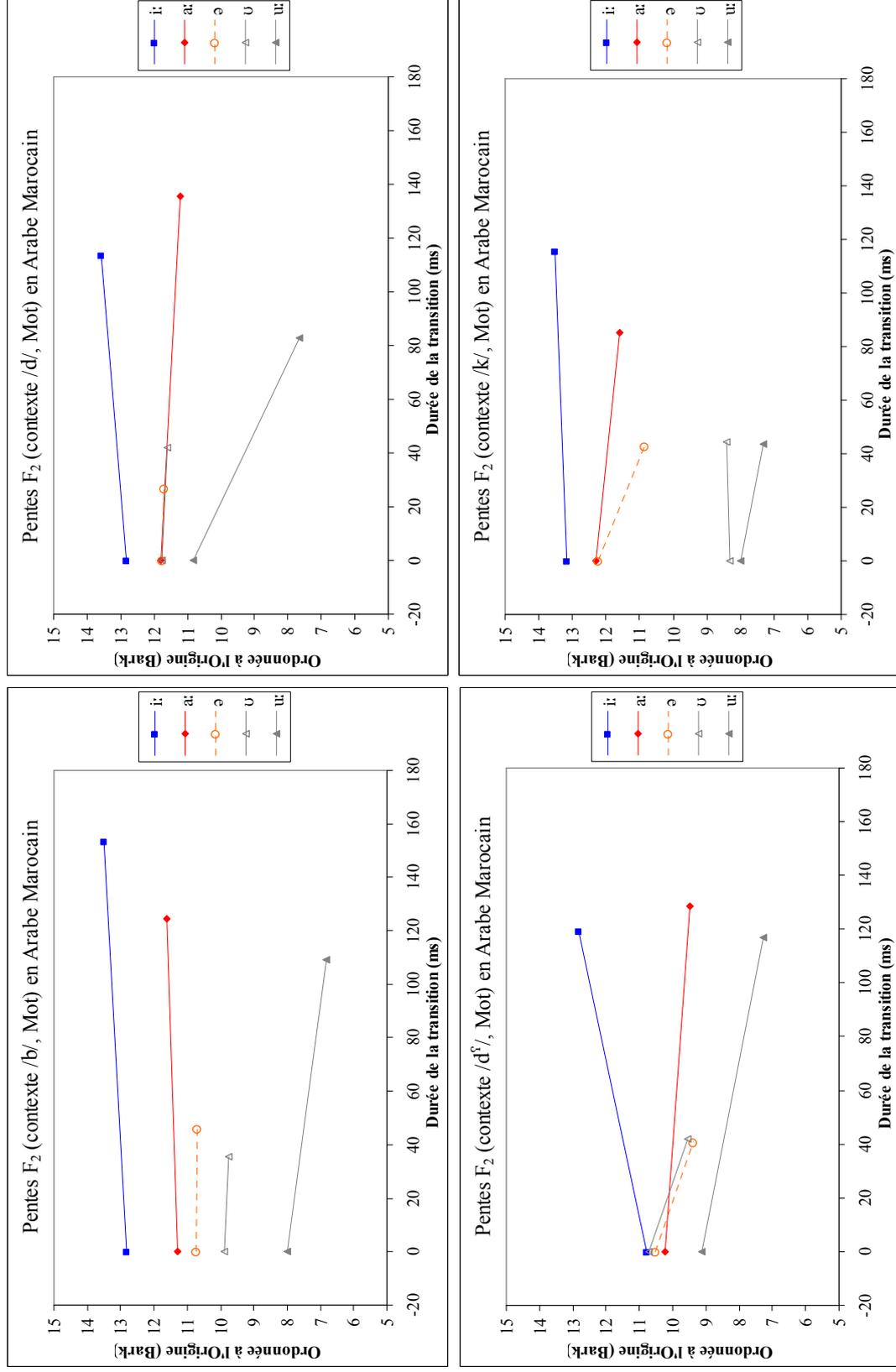


Figure 73 : Pentes formantiques de F₂ des voyelles en AM sans normalisation du temps, en fonction du lieu d'articulation en Mot.

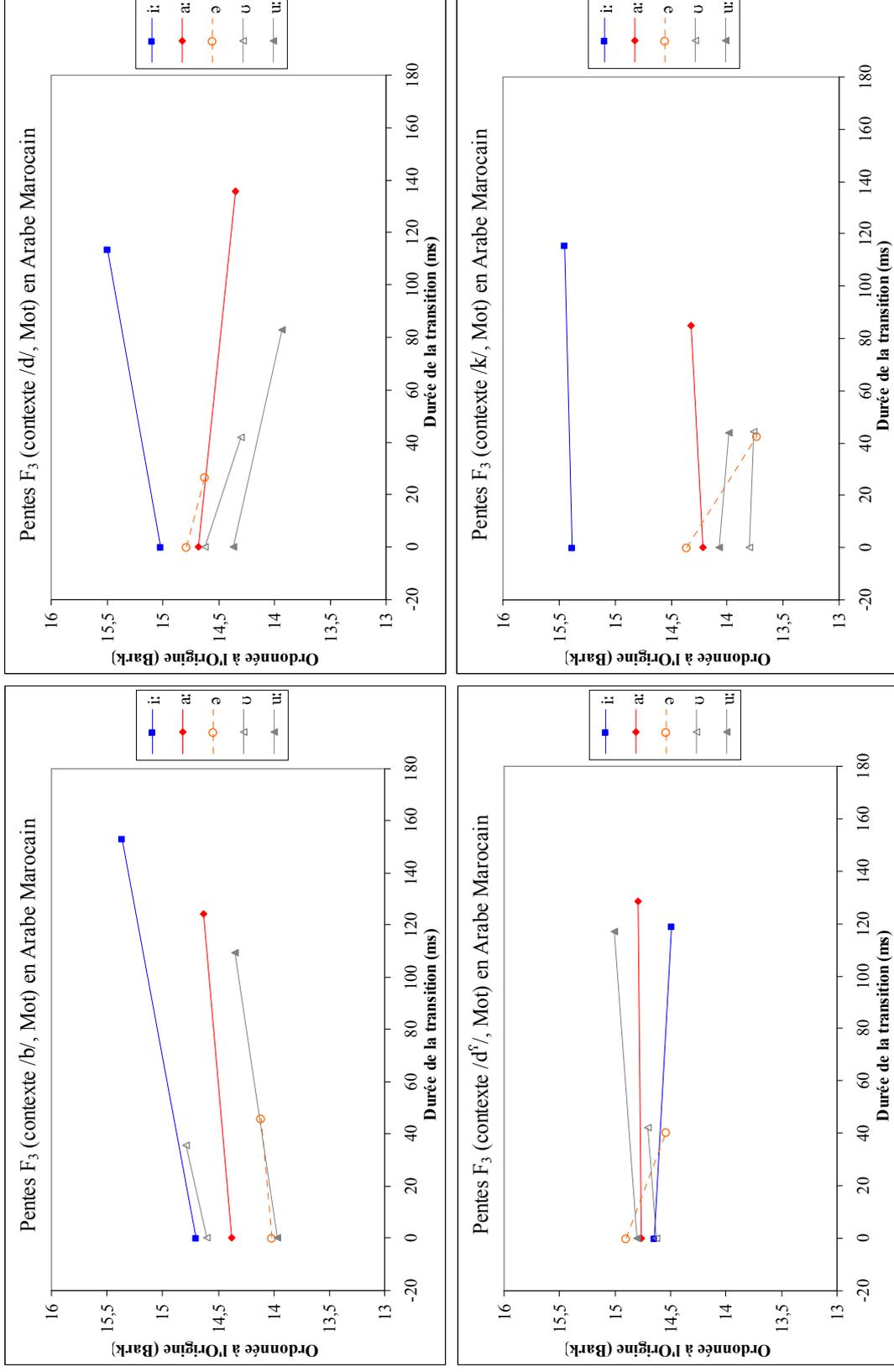


Figure 74 : Pentas formantiques de F₃ des voyelles en AM sans normalisation du temps, en fonction du lieu d'articulation en Mot.

Les résultats obtenus en ce qui concerne la production des voyelles en AM à partir des analyses **statiques**, montrent des différences importantes liées au timbre des voyelles brèves /ə u/ qui semblent « fusionnées » dans les environnements /d d^ʰ/, en revanche elles sont réalisées comme deux voyelles distinctes dans les environnements /b k/. En **dynamique**, les résultats montrent ces deux voyelles constituent une ou deux voyelles distinctes en fonction des valeurs des **pentés formantiques**, des points de début des transitions et des cibles vocaliques dérivées, en fonction des environnements consonantiques et les situations de production. Nous testerons dans la partie qui suit la validité de ces résultats avec l'aide de l'analyse discriminante qui nous permettra, une fois les confusions de classification examinées, de comprendre le statut des voyelles de l'AM. À l'issue de cette partie, nous présenterons une comparaison entre les taux de classification obtenus pour les différentes analyses discriminantes effectuées à partir des traits **statiques** vs **dynamiques**.

4.4.1.3. Résultats de l'analyse discriminante

4.4.1.3.1. En statique

Une analyse discriminante avec validation croisée, et ayant comme paramètres les valeurs formantiques de F₁, F₂ & F₃, obtenues à l'état stable, en incluant et en excluant la durée vocalique a été appliquée. Les taux de classification des voyelles de l'AM en fonction de la situation de production et du lieu d'articulation sont présentés dans le Tableau 19. Lorsque le paramètre de la durée vocalique est inclus dans l'analyse discriminante, une amélioration moyenne de 6% des taux de classification est observée (85,13% avec durée contre 79,32% sans durée), (BinomialQ ; p<0,001).

Tableau 19 : Taux de classification correcte des voyelles de l'AM à partir des valeurs formantiques de F₁, F₂ & F₃, en fonction de la situation de production et du lieu d'articulation. Les valeurs entre parenthèse indiquent les taux de classification en excluant la durée. (Taux significatifs : Box's M ; p<0,001).

	/b/	/d/	/d ^ʰ /	/k/
Mot	91,8% (82,7%)	94,8% (83,5%)	87,2% (80,4%)	87,7% (75,0%)
Syllabe	86,2% (83,3%)	80,3% (79,5%)	78,4% (75,2%)	93,9% (83,6%)
Isolation	81,3% (81,3%)	77,4% (76,6%)	71,7% (71,7%)	90,8% (79,0%)

Les taux de classification des voyelles de l'AM sont relativement élevés. La plupart des confusions dans les résultats de l'analyse discriminante sont explicables par le statut particulier des deux voyelles brèves /ə u/. Nous présentons sur les Figures 75 et 76 les taux des confusions de classification en situation Mot (pour les deux autres situations de production, voir Annexes 15 & 16). La proximité de ces deux voyelles a fait baisser les taux de classification des voyelles de l'AM produites dans les trois environnements /b d d^h/ ; dans l'environnement /k/ en revanche, le taux de classification correcte de ces deux voyelles est en moyenne proches des 85~90%. Dans les trois situations de production, les trois voyelles longues /i: a: u:/ ont été classifiées la plupart du temps avec des taux proches des 95%. Lorsque le paramètre de la durée est exclu de l'analyse discriminante, les confusions dans la classification des voyelles sont plus importantes.

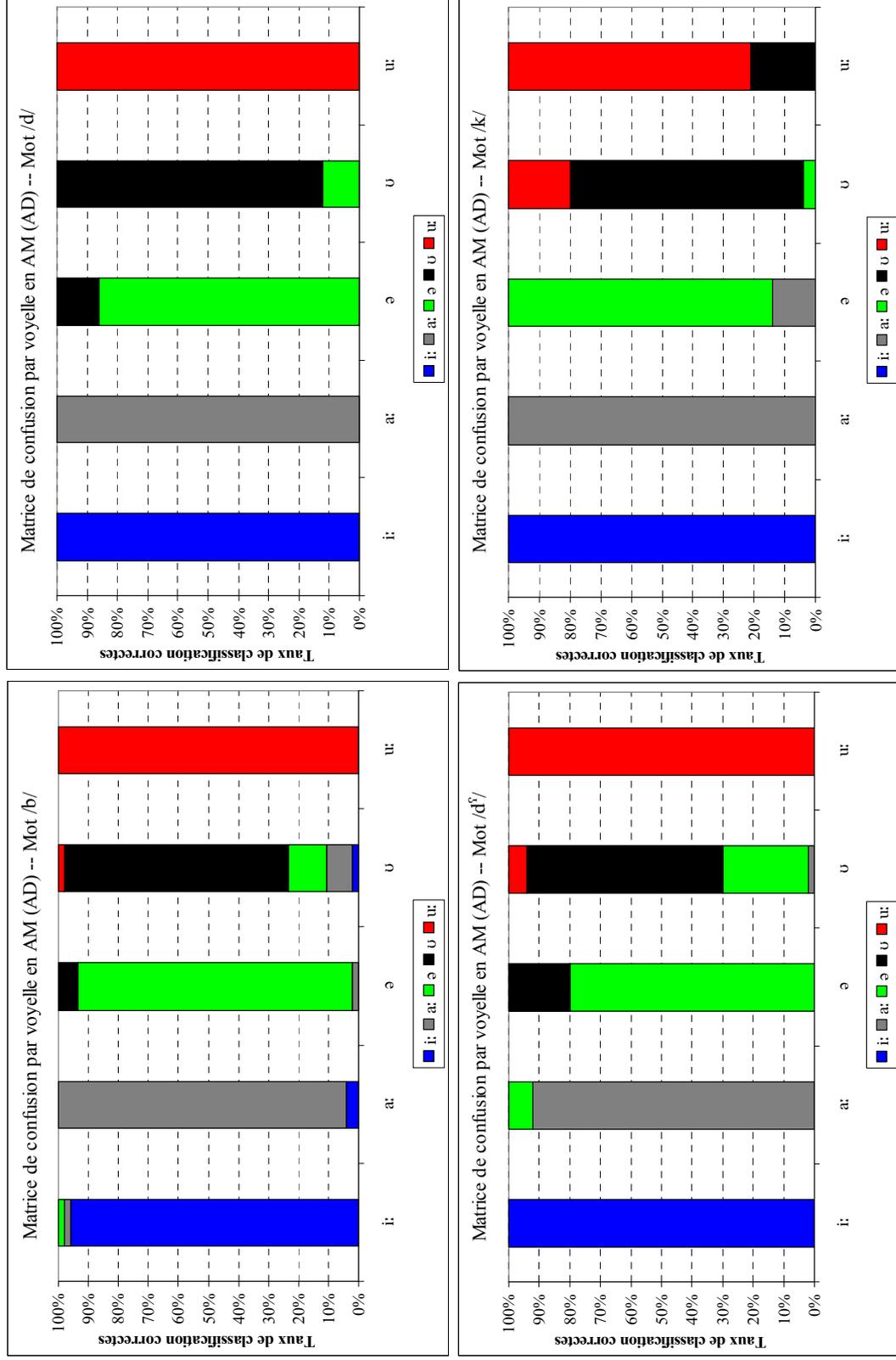


Figure 75 : Matrices de confusion de classification des voyelles de l'AM en **statique** avec la durée vocalique (AD) en situation de production Mot et en fonction du lieu d'articulation.

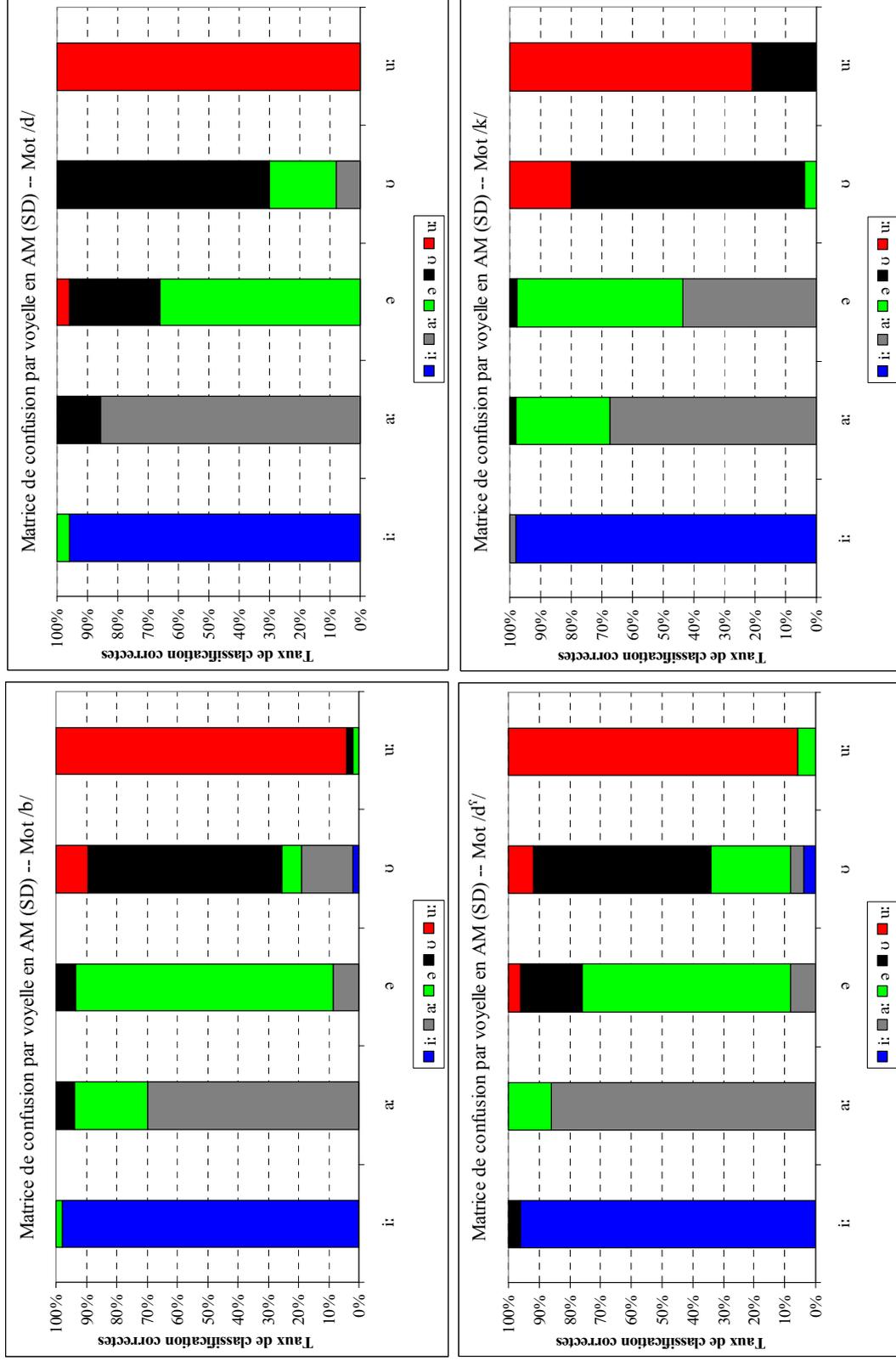


Figure 76 : Matrices de confusion de classification des voyelles de l'AM en **statique** sans la durée vocalique (SD) en situation de production Mot et en fonction du lieu d'articulation.

4.4.1.3.2. En dynamique

4.4.1.3.2.1. Régression linéaire

Les paramètres utilisés dans cette analyse sont les valeurs des **pentés formantiques** et des ordonnées à l'origine de F_1 , F_2 & F_3 des voyelles de l'AM, ainsi que la durée de la transition formantique. Les taux de classification des voyelles de l'AM en fonction de chaque situation de production et par lieu d'articulation sont présentés dans le Tableau 20.

Tableau 20 : Taux de classification correcte des voyelles de l'AM à partir de la régression linéaire de F_1 , F_2 & F_3 en fonction de la situation de production et du lieu d'articulation. (Taux significatifs : Box's M ; $p < 0,001$).

	/b/	/d/	/dʰ/	/k/
Mot	91,2%	88,3%	76,0%	87,2%
Syllabe	82,9%	82,3%	79,6%	91,8%
Isolation	79,6%	78,7%	75,9%	90,8%

Nous obtenons une légère amélioration des taux de classification par rapport à ceux obtenus en **statique**. Comme avec l'analyse discriminante en **statique**, les taux des classifications variables sont dus aux rapprochements entre les deux voyelles /ə ʊ/, qui occasionnent des confusions de classification importantes. Sur la Figure 77 les confusions de classification des voyelles de l'AM en situation Mot sont présentées (pour les confusions en situations Syllabe et Isolation, voir Annexe 17). Les trois voyelles longues /i: a: u:/ sont toujours bien classifiées avec des taux proches des 100% ; les deux voyelles brèves /ə ʊ/ obtiennent des taux variables. La voyelle /ə/ est classifiée correctement avec un taux dépassant les 90% seulement dans l'environnement /b/ ; dans les autres environnements elle est souvent confondue avec la voyelle /ʊ/. Dans l'environnement /k/, la voyelle /ʊ/ est confondue seulement avec /u:/ et pas du tout avec la voyelle /ə/. Dans les autres situations de production, nous obtenons approximativement les mêmes résultats.

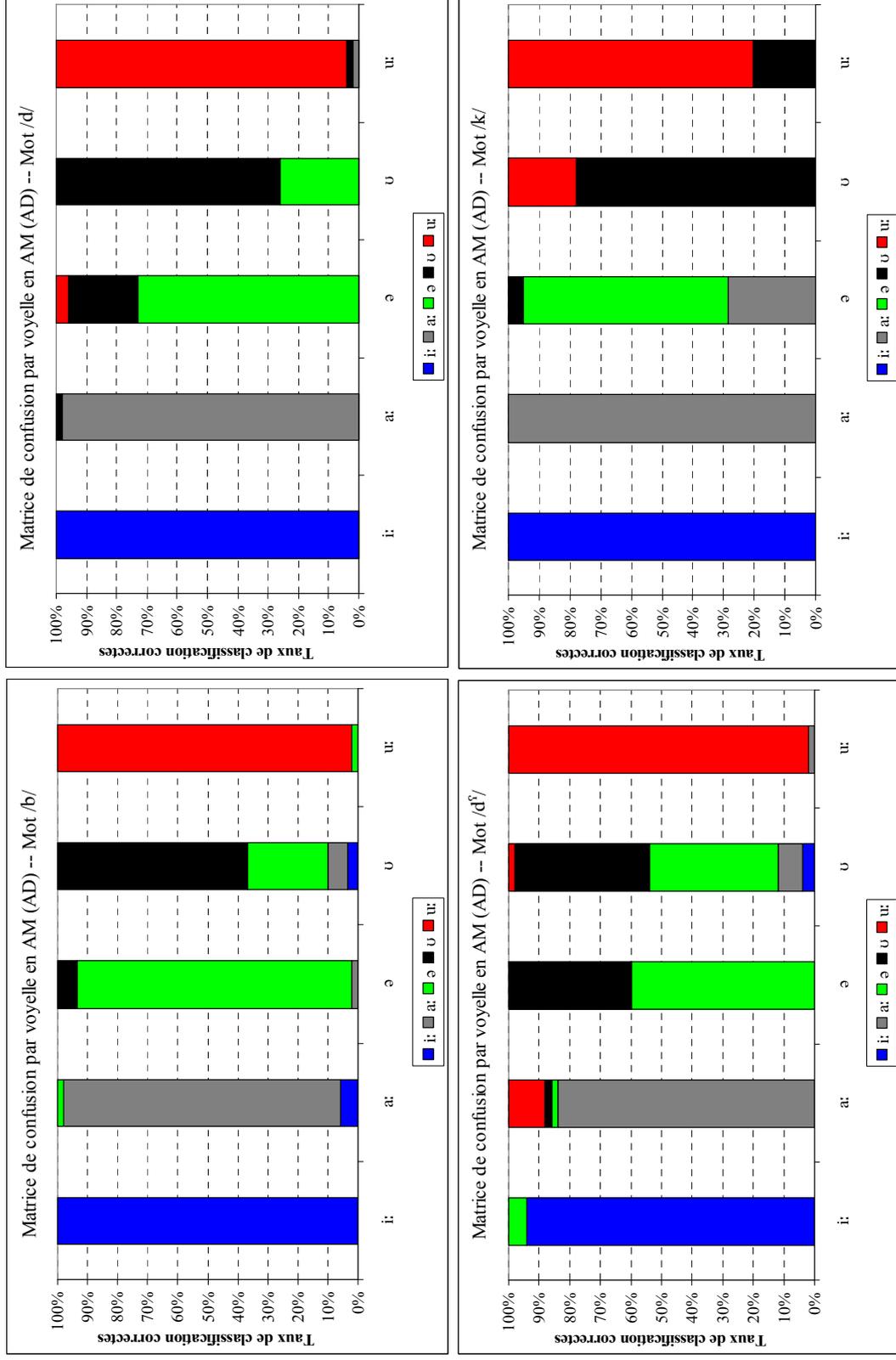


Figure 77 : Taux de classification des voyelles de l'AM en dynamique (régression linéaire) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

4.4.1.3.2.2. Régressions polynomiales de 2^{ème} et 3^{ème} ordre

Les paramètres utilisés dans cette analyse sont les coefficients des deux analyses de régressions polynomiales de 2^{ème} et 3^{ème} ordre⁶³, obtenus pour F₁, F₂ & F₃ des voyelles de l'AM, ainsi que la durée de la transition formantique. Les taux de classification correcte des voyelles en fonction de chaque situation de production et de chaque lieu d'articulation sont présentés dans le Tableau 21. Nous obtenons des taux de classification relativement élevés dans les deux analyses. À partir des taux de classification correcte obtenus avec les deux analyses de régression polynomiale, nous n'obtenons aucune amélioration, en moyenne, du taux de classification des voyelles en AM en augmentant l'ordre des polynômes (un taux de classification correcte moyen de 83,2% avec les deux analyses). Dans certains contextes consonantiques, la régression polynomiale de 3^{ème} ordre améliore les taux de classification (e.g., contexte /b/ en situation Mot) et dans d'autres, elle baisse les taux (e.g., contexte /d^s/ en situation Syllabe). Il semblerait que pour certaines trajectoires formantiques, l'augmentation de l'ordre de l'analyse de régression est pertinente et pour d'autre, une simple analyse linéaire est suffisante.

Tableau 21 : Taux de classification correcte des voyelles de l'AM à partir des coefficients des régressions polynomiales de 2^{ème} et 3^{ème} ordre obtenus pour F₁, F₂ & F₃, en fonction de la situation de production et du lieu d'articulation. La valeur entre parenthèses indique le taux de classification en 3^{ème} ordre. (Taux significatifs : Box's M ; p<0,001).

	/b/	/d/	/d ^s /	/k/
Mot	88,6% (90,9%)	87,7% (88,5%)	80,6% (81,5%)	84,4% (82,6%)
Syllabe	81,7% (82,5%)	82,3% (80,7%)	77,2% (72,0%)	91,0% (91,4%)
Isolation	79,6% (79,6%)	79,5% (82,3%)	74,7% (76,3%)	90,8% (89,5%)

Comme avec la régression linéaire, la plupart des confusions sont dues au rapprochement entre les deux voyelles brèves en AM. Nous présentons sur les Figures 78 et 79 les taux de confusion dans la classification des voyelles de l'AM produites en situation

⁶³ Les coefficients utilisés sont : pour la régression polynomiale de 2^{ème} ordre, a₀ = la moyenne des valeurs, a₁ = la droite de régression et a₂ = l'ordonnée à l'origine ; pour la régression polynomiale de 3^{ème} ordre, a₀ = la moyenne des valeurs, a₁ = la droite de régression, a₂ = l'excursion parabolique de la voyelle indépendamment de sa droite et a₃ = l'ordonnée à l'origine

Mot, avec les régressions polynomiales de 2^{ème} et 3^{ème} ordre, respectivement (pour les deux autres situations de production, voir Annexes 18 & 19). Les trois voyelles longues /i: a: u:/ obtiennent les taux de classification les plus élevés, tandis que les brèves /ə ʊ/ sont souvent confondues entre elles dans les environnements /b d d^f/ ; dans l'environnement /k/, la voyelle /ə/ a été mal classifiée comme /a: ʊ/, en revanche, la voyelle /ʊ/ a été mal classifiée comme /u:/ et jamais comme /ə/. Les taux de classification obtenus dans les deux autres situations de production sont homogènes avec ceux en Mot.

Nous observons que lorsque l'ordre de la régression polynomiale est augmenté de 2^{ème} au 3^{ème} ordre, les confusions dans la classification des voyelles augmentent. Comme nous l'avons déjà vu avec la comparaison inter-langues, l'augmentation de l'ordre de la régression polynomiale semble ne pas être utile pour l'analyse des monophthongues (McDougall, communication personnelle).

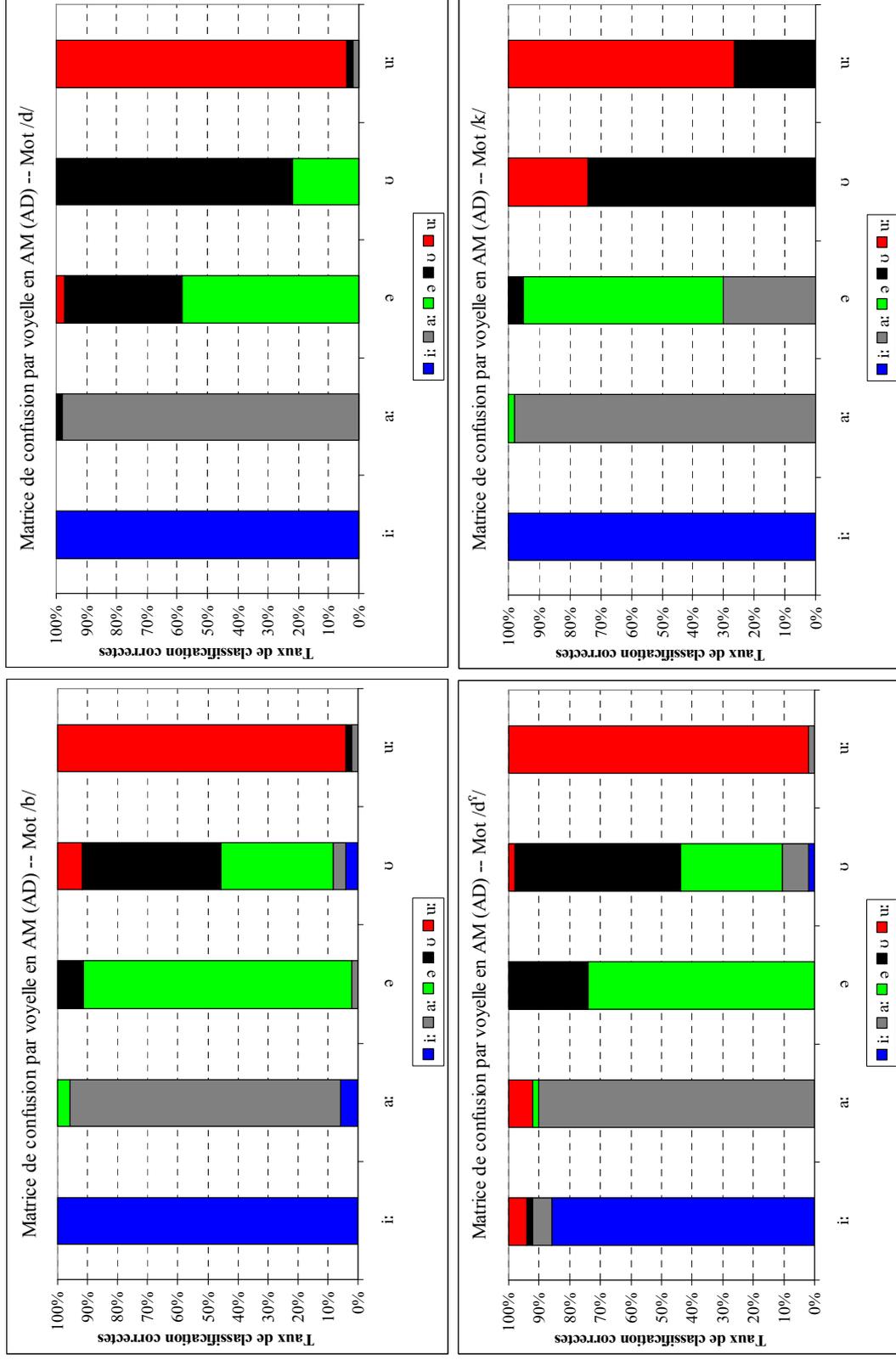


Figure 78 : Matrices de confusion de classification des voyelles de l'AM en **dynamique** (régression polynomiale de 2^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

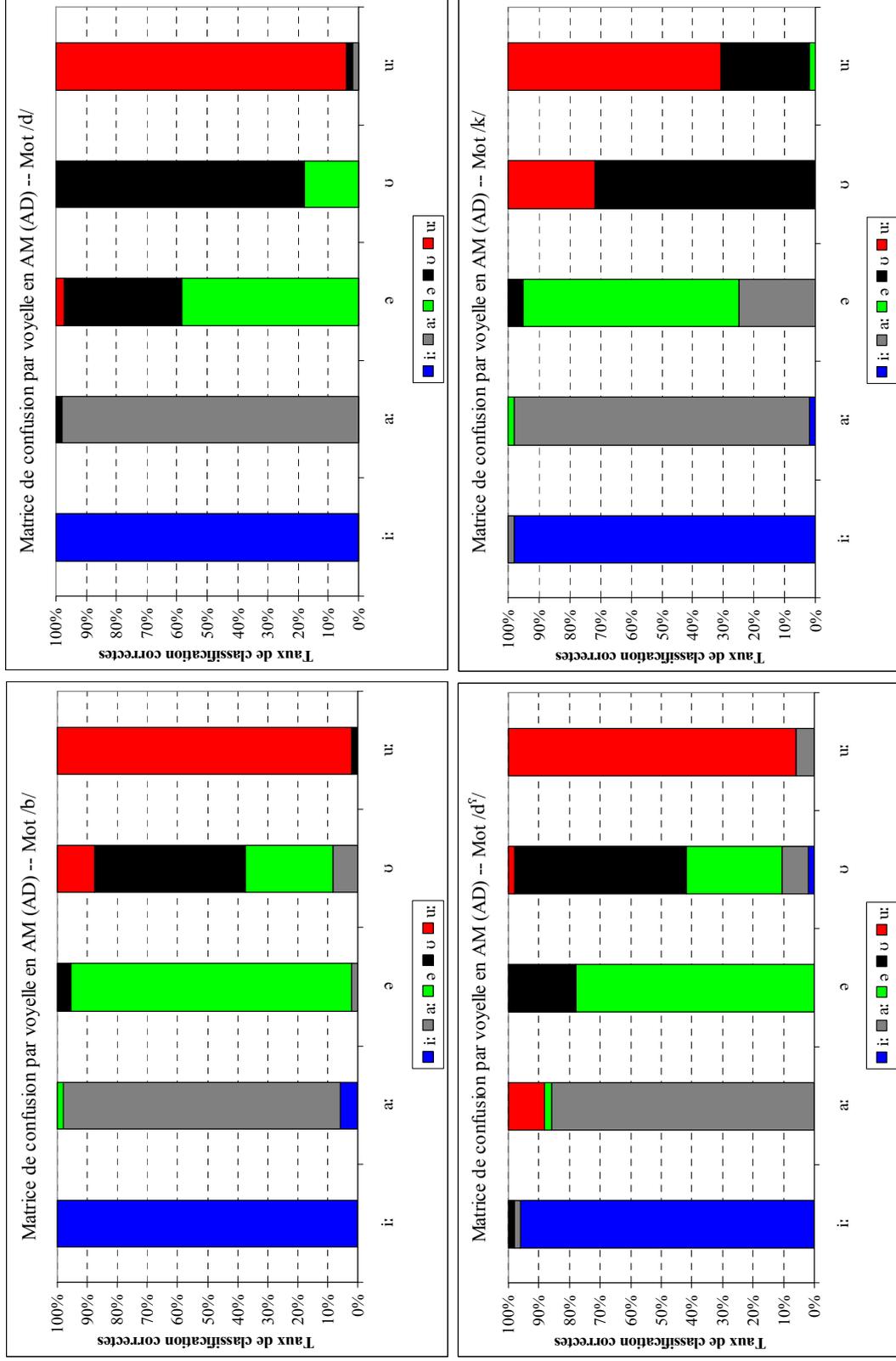


Figure 79 : Matrices de confusion de classification des voyelles de l'AM en **dynamique** (régression polynomiale de 3^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

4.4.1.4. Discussion des résultats en statique vs en dynamique

Les résultats obtenus dans les précédentes parties et plus particulièrement les taux d'erreurs de classification des voyelles de l'AM obtenus en **statique** (à partir des « cibles vocaliques ») et en **dynamique** (à partir des coefficients des régressions linéaire et polynomiale de 2^{ème} et 3^{ème} ordre) peuvent être corrélés aux différences inter-individuelles. En effet, nous avons vu que la taille des ellipses de dispersion des voyelles (présentée sur la Figure 68, p. 213), mais également les différences de « cibles vocaliques » produites par les locuteurs marocains (voir Figures 69 à 71), sont relativement grandes, indiquant que les stratégies individuelles lors de la réalisation des voyelles sont différentes. L'analyse discriminante a permis de discriminer les voyelles produites en AM et plus particulièrement les voyelles longues ; seules les deux voyelles brèves /ə u/ ont été souvent confondues soit entre elles, soit c'est /u:/ qui est confondue avec /u/ (dans l'environnement /k/ seulement). L'utilisation des différentes analyses **statiques** et **dynamiques** a permis de mettre en évidence les différences existantes entre les voyelles en AM. Il semblerait toutefois que la **dynamique** permet de mieux rendre compte des variations vocaliques et des trajectoires formantiques **intrinsèques** et **extrinsèques** des voyelles de l'AM ; constat obtenu à partir de la comparaison des taux de classification des différentes analyses discriminantes (voir Figure 80). Lorsque les taux de classification moyens obtenus à partir des analyses **statiques** sont comparés à ceux obtenus à partir de l'une des trois analyses de régression, les résultats montrent une augmentation significative de 4% en **dynamique** quels que soit la situation de production et le lieu d'articulation (BinomialQ, $p < 0,001$). Lorsque la durée vocalique est ajoutée aux valeurs formantiques de F_1 , F_2 & F_3 , une augmentation significative des taux de classification est obtenue. En effet, en situation Mot, le taux de classification obtenu, avec les paramètres **statique** et durée vocalique, augmente significativement de 10% par rapport aux paramètre **statique** sans durée (BinomialQ ; $p < 0,001$) et de 4% par rapport aux paramètres **dynamiques** (obtenus par les trois analyses de régressions), (BinomialQ ; $p < 0,001$). En situation Syllabe, le taux augmente significativement de 5% par rapport aux taux en **statique** (BinomialQ ; $p < 0,001$) ; de 3% avec les paramètres de la régression polynomiale de 2^{ème} ordre

(BinomialQ ; $p < 0,001$) et de 2% avec les paramètres de la régression polynomiale de 3^{ème} ordre (BinomialQ ; $p < 0,001$) ; aucune différence n'est obtenue avec la régression linéaire. En situation Isolation enfin, la tendance s'inverse, car une augmentation significative de 2% est obtenue entre la classification avec les paramètres de la régression polynomiale de 3^{ème} ordre et celle avec les paramètres **statique** & durée vocalique (BinomialQ ; $p < 0,05$).

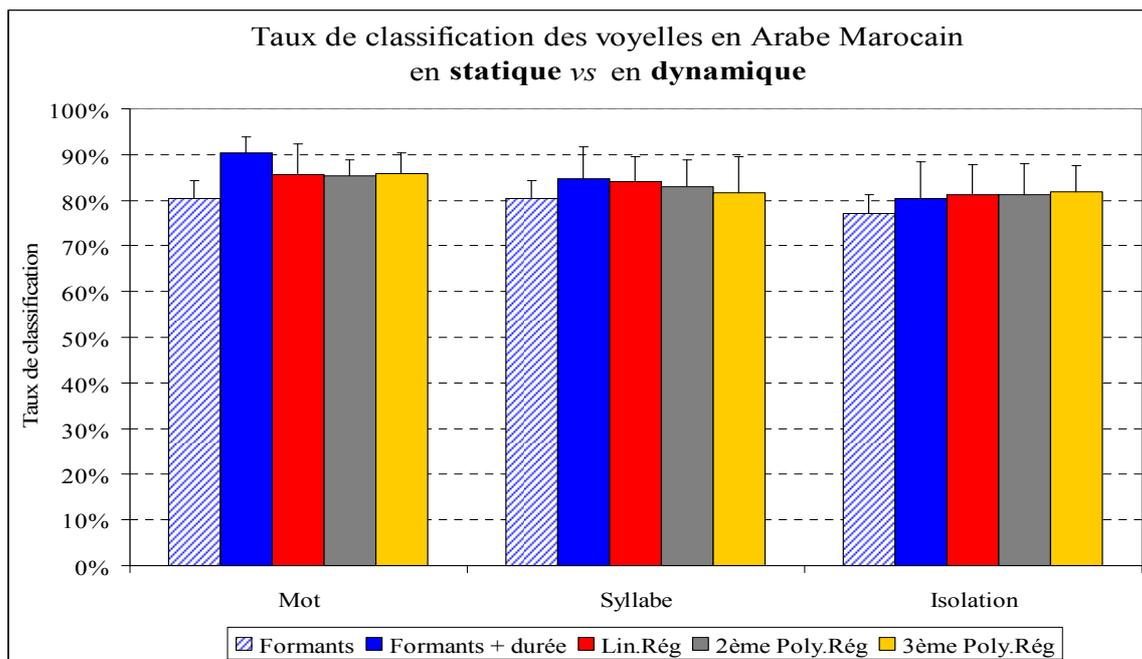


Figure 80 : Taux de classification correcte moyens des voyelles de l'AM produites dans les contextes /b d d^f k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur). Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} ou 3^{ème} ordre).

Les résultats de l'analyse discriminante laisse supposer que la durée vocalique semble être un « indice » primordial qui a permis la discrimination entre les voyelles de l'AM plus particulièrement en situation de production Mot et Syllabe. Dans la situation Isolation, les taux ont augmenté significativement avec les autres analyses **dynamiques** indiquant ainsi que les trajectoires formantiques **intrinsèques** des voyelles de l'AM ont permis de les discriminer. Nous nous attendons à ce que ces différents paramètres « **dynamiques** », i.e., la durée vocalique et les trajectoires formantiques propres aux voyelles permettent aux auditeurs marocains de discriminer les voyelles de leur système.

En ce qui concerne le système vocalique de l'AM, la littérature explique qu'il est constitué de trois voyelles longues /i: a: u:/ auxquelles vient s'ajouter une ou deux voyelles brèves /ə/ et /ʊ/ : la première étant la résultante de la « fusion » des deux anciennes voyelles

brèves /i a/ héritées de l'arabe classique ; tandis que la seconde est la variante brève de la voyelle /u/. Dans certains travaux, on décrit ce système comme n'ayant qu'une seule voyelle brève /ə/, tandis que pour d'autres, ce sont deux voyelles brèves qui s'opposent (voir § 1.4.2). Nos résultats ont montré que les locuteurs marocains natifs ont produit deux voyelles brèves distinctes dans les trois situations de production en Mot, Syllabe et Isolation. Les taux de classification élevés de la voyelle /u/ dans tous les environnements consonantiques (et plus particulièrement dans l'environnement /k/) nous laissent penser qu'il pourrait s'agir de deux voyelles différentes qualitativement. Pourrions-nous alors les considérer comme deux variations phonétiques d'un seul timbre vocalique /ə/ ou comme deux réalisations de deux timbres différents /ə u/ ? À partir de notre corpus, nous pouvons conclure qu'il pourrait s'agir de deux timbres vocaliques brefs différents, ainsi le système vocalique de l'AM que nous proposons est constitué des cinq voyelles /i: a: ə u u/. Nous testerons la validité de notre hypothèse en perception des voyelles (voir Chapitre 5), à partir de laquelle nous identifierons les attentes perceptives des auditeurs marocains natifs.

Intéressons-nous maintenant aux spécificités du système vocalique de l'AJ.

4.4.2. *En Arabe Jordanien*

Nous nous intéressons dans cette partie aux spécificités du système vocalique de l'AJ et plus particulièrement au statut des deux voyelles brèves /i u/. Nous avons montré dans une précédente étude que ces deux voyelles sont produites comme /e o/ respectivement dans les environnements /d d^ʕ/ (voir Al-Tamimi, 2001, 2002 ; Al-Tamimi & Barkat-Defradas, 2002 ; Barkat-Defradas *et al.*, 2003). Nous étendons notre étude à d'autres environnements consonantiques /b k/ (en plus de /d d^ʕ/) et situations de production (en Mot, Syllabe et Isolation). Dans la première partie, nous présentons les résultats des « cibles vocaliques » (en **statique**) produites par les locuteurs jordaniens et ensuite, les résultats des **pentés formantiques** (en **dynamique**) obtenues à partir de différentes analyses de régression (linéaire et polynomiales). Nous terminons cette partie par une présentation des résultats des analyses discriminantes qui nous permettront de valider nos résultats et de discriminer les voyelles de l'AJ. Commençons avec les résultats en **statique**.

4.4.2.1. Résultats en statique

Les résultats graphiques présentés sur la Figure 81 montrent que les espaces vocaliques de l'AJ obtenus dans les trois situations de production sont significativement non différents ni sur F_1 ($F(2, 4190) = 0,53$; $p = 0,59$), ni sur F_2 ($F(2, 4190) = 0,53$; $p = 0,59$). Ces résultats vont dans le sens des difficultés rencontrées par les locuteurs jordaniens. Comme nous l'avons expliqué auparavant, les locuteurs jordaniens ont rencontré de grandes difficultés lors de la réalisation de cette tâche. En effet, il leur a été difficile de produire les voyelles de leur système dans les deux situations de production Syllabe et Isolation, car c'était une tâche très artificielle pour eux. Ainsi, pour certains parmi eux, ils produisaient toutes les voyelles en situation Isolation soit comme une voyelle [ə] soit comme une voyelle [a:]. Les aires de dispersion des voyelles dans les trois situations de productions sont significativement non différentes : 9,51 Bark² en Mot ; 8,48 Bark² en Syllabe et 10,07 en Isolation ($F(2, 108) = 2,12$; $p = 0,13$).

Nos résultats ne vont pas dans le sens de l'hypothèse selon laquelle l'espace vocalique obtenu en isolation est plus grand que celui en Mot (Lindblom, 1990). En revanche, l'interaction entre la situation de production et la voyelle s'est révélée significative sur les deux axes, mais elle n'est pas toujours en faveur de l'agrandissement de l'espace vocalique en Isolation : sur F_1 , les voyelles /i: i e: u o:/ sont plus « fermées » en Isolation et plus « ouvertes » en Mot ; /a/ est plus « fermée » en Mot et plus « ouverte » en Isolation et le même degré d'aperture est obtenu pour /a:/ dans ces deux situations de production ($F(14, 4190) = 7,88$; $p < 0,001$) ; sur F_2 , en revanche, les voyelles /i: i e:/ sont plus « antériorisées » en Isolation et plus « postériorisées » en Mot ; et les voyelles /a: a o: u u:/ sont plus « postériorisées » en Isolation et plus « antériorisées » en Mot, ($F(14, 4190) = 7,86$; $p < 0,001$).

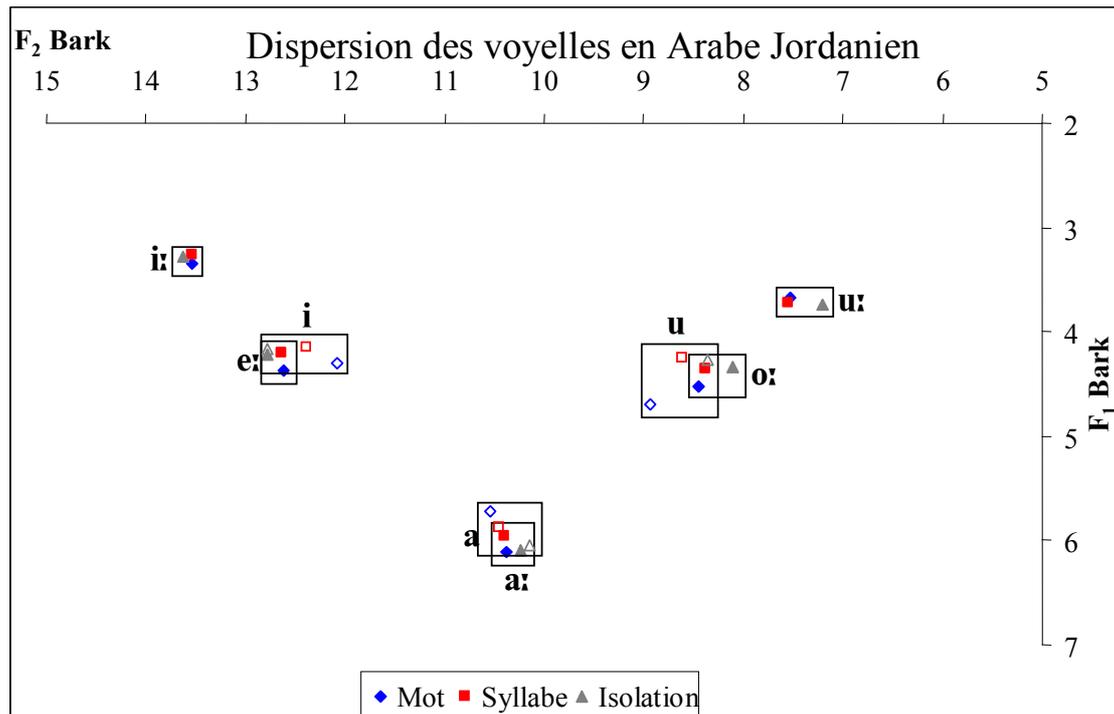


Figure 81 : Espace vocalique en AJ en fonction de la situation de production. Les voyelles brèves sont représentées par les symboles avec les fonds vides.

Bien que l'espace vocalique obtenu en fonction des situations de production ne soit pas significativement plus « périphérique » en Isolation par rapport à celui en Mot ; les résultats statistiques de l'interaction entre la situation de production et le lieu d'articulation montre des différences significatives (voir Figure 82) : sur F_1 , et dans les environnements /b d^s/, c'est dans la situation de production Mot que l'espace est significativement plus ouvert ; tandis que pour les environnements /d k/, ce sont les situations Isolation et Syllabe qui présentent les espaces les plus ouverts ($F(6, 4190) = 2,99$; $p < 0,01$) ; sur F_2 , et dans les environnements /b d^s/, c'est en Isolation que l'espace est significativement plus antérieur ; tandis que pour les environnements /d k/, c'est la situation Syllabe qui présente l'espace le plus antérieur ($F(6, 4190) = 10,68$; $p < 0,001$).

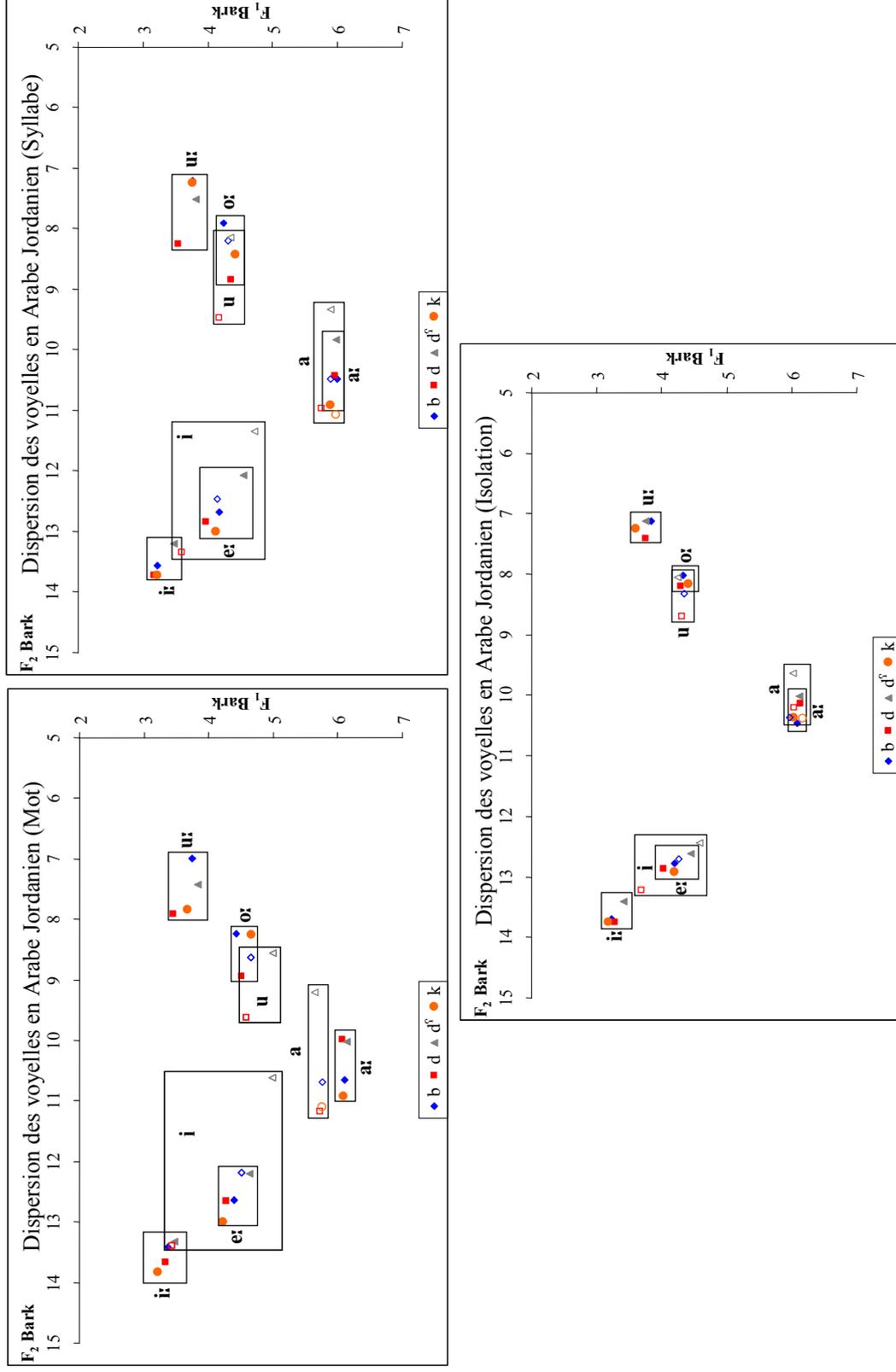


Figure 82 : Espaces vocaliques en AJ en fonction de chaque situation de production et du lieu d'articulation des consonnes /b d d' k/. Les voyelles brèves sont représentées par les symboles avec les fonds vides.

Les aires de dispersion des voyelles produites dans les trois situations de production en fonction de chaque lieu d'articulation sont présentées sur la Figure 83. Les résultats graphiques donnent une tendance à l'agrandissement de la taille de l'aire en Isolation dans les environnements consonantiques /b d k/ par rapport aux deux situations Mot et Syllabe (effet non significatif ($F(6, 108)=0,32 ; p=0,93$)).

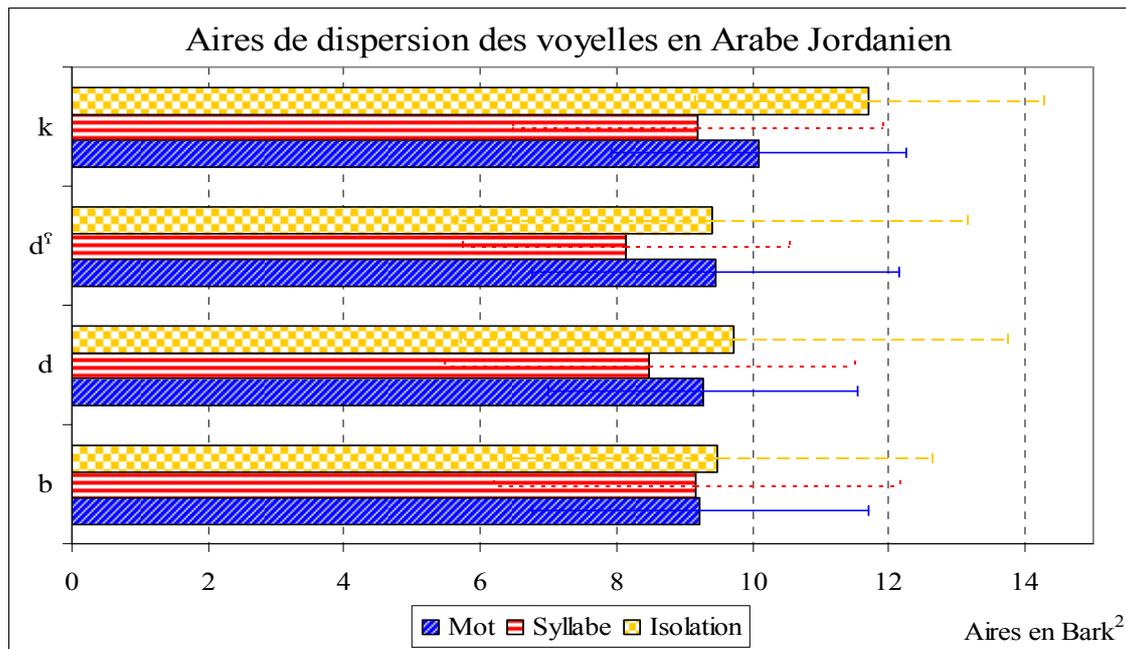


Figure 83 : Aires de dispersion (moyenne et écart-type) des voyelles de l'AJ en fonction de la situation de production et du lieu d'articulation.

La comparaison des timbres vocaliques obtenus en fonction de chaque situation de production et chaque lieu d'articulation donne des différences significatives seulement sur F_2 : dans les trois situations de production, les voyelles antérieures produites dans l'environnement /k/ sont plus « antériorisées » que dans les autres environnements consonantiques ; les voyelles postérieures sont plus « postériorisées » dans les environnements /dʳ k/, ($F(42, 4190) = 3,82 ; p < 0,001$). L'étude des voyelles problématiques ; i.e., les voyelles brèves /i a u/ et leurs statuts par rapport à leurs correspondantes longues /i: a: u:/, les résultats graphiques (voir Figure 82) et statistiques montrent que la voyelle /i/ est significativement différente de sa correspondante longue /i:/ sur les deux axes : sur F_1 , la voyelle /i/ est significativement plus « ouverte » que /i:/ et se rapproche de la voyelle longue /e:/, quels que soient la situation de production et le lieu d'articulation (sauf dans

l'environnement /d/ où l'influence de l'entourage contextuel ([dijja]) a affecté le degré d'aperture de cette voyelle) ($F(2, 1592) = 556,45$; $p < 0,001$) ; sur F_2 , /i/ est significativement plus « postérieure » que /i: e:/, toujours quels que soient la situation de production et le lieu d'articulation ses fréquences formantiques se rapprochent néanmoins davantage que de celles de la voyelle /e:/, ($F(2, 1592) = 302,22$; $p < 0,001$). En ce qui concerne la voyelle /u/, elle est plus « ouverte » sur F_1 ($F(2, 1439) = 30,28$; $p < 0,001$) et plus « antériorisée » sur F_2 ($F(2, 1439) = 84,54$; $p < 0,001$) que sa correspondante longue /u:/ quels que soient la situation de production et le lieu d'articulation ; elle se rapproche de la voyelle longue /o:/. Enfin, la voyelle /a/ est significativement plus « ouverte » sur F_1 que sa correspondante longue ($F(1, 1159) = 29,67$; $p < 0,001$) dans la situation de production Mot seulement et quelque soit le lieu d'articulation ; dans les deux autres situations de production et pour les trois lieux d'articulation, aucune différence n'est obtenue ($F(6, 1159) = 0,45$; $p = 0,85$) ; aucun effet n'est observé sur F_2 . Ces résultats montrent que les trois voyelles brèves de l'AJ ont changé de timbre : la voyelle fermée antérieure /i/ s'est transformée en voyelle mi-fermée antérieure /e/ ; /u/ est devenue une voyelle mi-fermée postérieure /o/ ; et enfin la voyelle /a/ conserve la qualité de son timbre.

Nous pouvons nous poser la question suivante : est-ce que le changement de degré d'aperture obtenu pour les deux voyelles /i u/ n'est pas le résultat des structures syllabiques utilisées ? On sait que lorsque les voyelles fermées en arabe et plus particulièrement en AJ sont produites dans une structure syllabique de type CVC, celles-ci se trouvent plus ouvertes (Ghazali, 1979 ; Bani-Yassin & Owens, 1987 ; Barkat, 2000 ; etc...). Dans notre étude, cet effet est avéré dans la situation de production Mot, seulement dans le cas de la voyelle /a/ qui est réalisée comme une voyelle plus « fermée » que sa correspondante longue. En ce qui concerne les situations Syllabe et Isolation, la réalisation de la voyelle /a/ n'a pas été altérée, i.e., la qualité de son timbre est identique à celle de sa correspondante longue. En revanche, pour ce qui est des deux voyelles brèves /i u/, les résultats dans les deux situations Syllabe et Isolation montrent qu'elles sont réalisées comme des voyelles mi-fermées. Si la réalisation de ces deux voyelles était influencée par la structure syllabique CVC, alors dans ce cas, leurs productions dans les situations CV ou V devaient donner des voyelles fermées significativement différentes. Or, nous obtenons dans tous les cas de figures, des voyelles mi-

fermées (à part la voyelle /i/ dans l'environnement /d/ qui est réalisée comme une voyelle fermée : effet explicable par l'entourage contextuel). Nous reviendrons sur cette question dans la discussion générale.

Nous avons voulu examiner les effets de la situation de production et des lieux d'articulation sur la taille des ellipses de dispersion de chaque voyelle (voir Figure 84 et Tableau 22). Nous observons, en moyenne, une aire de dispersion des voyelles moins importante en Mot et plus importante en Isolation et/ou Syllabe (sauf pour la voyelle /a:/).

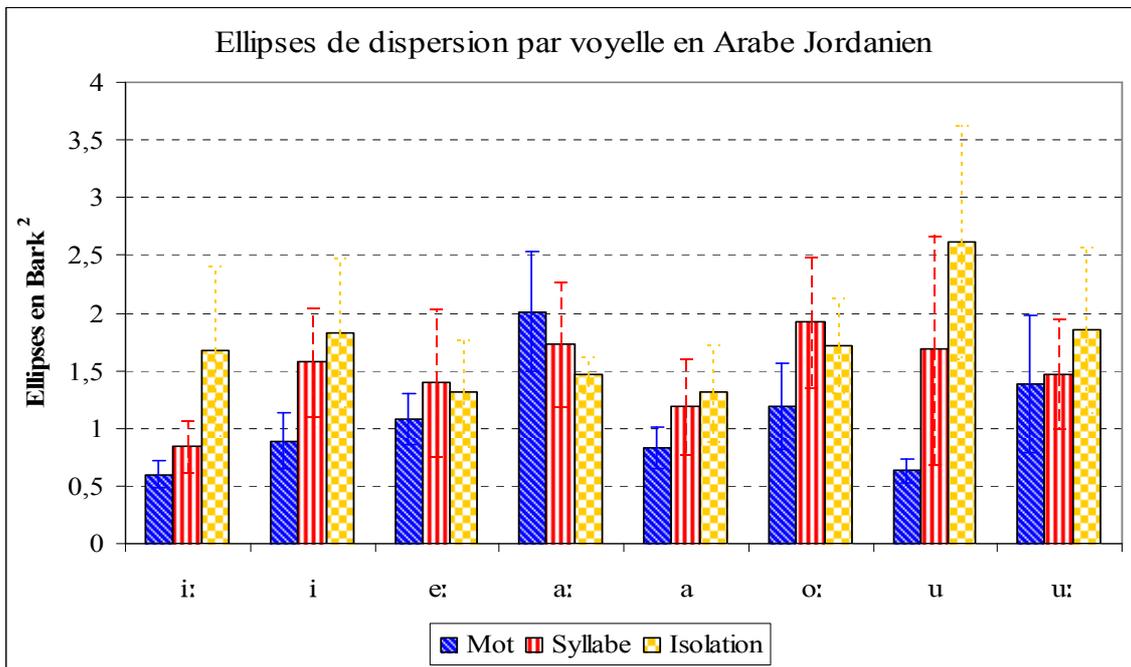


Figure 84 : Ellipses de dispersion (moyenne et écart-type) des voyelles en fonction de la situation de production en AJ.

Tableau 22 : Taille des ellipses de dispersion des catégories vocaliques en fonction de la situation de production et du lieu d'articulation en AJ.

(Bark ²)		i:	i	e:	a:	a	o:	u	u:
Mot	b	0,53	1,17	0,77	1,94	1,05	1,28	0,73	1,51
	d	0,75	0,73	1,19	1,79	0,93	1,50	0,52	0,83
	d^ʰ	0,64	0,78	1,06	0,69	2,75		0,65	1,02
	k	0,49		1,29	1,57	0,67	0,77		2,17
Syllabe	b	0,80	1,14	0,71	1,45	1,48	1,31	2,82	1,84
	d	0,68	1,51	1,02	1,26	1,04	2,43	1,26	1,47
	d^ʰ	1,16	2,07	2,02	2,48	0,68		0,98	0,79
	k	0,74		1,85	1,74	1,57	2,04		1,79
Isolation	b	1,19	1,14	0,85	1,56	1,10	1,83	3,73	2,59
	d	1,60	2,42	1,08	1,32	1,93	2,06	2,35	2,35
	d^ʰ	2,74	1,93	1,36	1,64	1,17		1,77	1,28
	k	1,19		1,94	1,35	1,03	1,28		1,20

Certains des résultats obtenus en AJ ne vont pas dans le sens des propositions de la littérature : la taille des aires de dispersion intra-catégorie vocalique est grande en contexte Isolation et plus petite en contexte Mot. Or d'après les propositions de la théorie *H&H*, la précision articulatoire conduit à une réduction de la dispersion des catégories vocaliques en contexte Isolation. Comme nous l'avons déjà expliqué, ces résultats peuvent être corrélés aux difficultés rencontrées par les locuteurs jordaniens : en Syllabe, les locuteurs jordaniens avaient tendance à produire des voyelles extrêmes, tandis que lorsqu'ils produisaient les voyelles en contexte Isolation, le timbre de ces voyelles correspondait à une production *Hypo-Articulée* tendant vers une voyelle centrale. Nous observons une variabilité inter-individuelle importante quantifiée par des écart-types importants et des tailles des aires de dispersion intra-catégorie vocalique très grandes. Cette variabilité est à l'origine d'une dispersion différente des voyelles en AJ en fonction de la situation de production et du lieu d'articulation. Nous présentons ainsi sur les Figures 85 à 87, la dispersion des voyelles en AJ à partir des données individuelles.

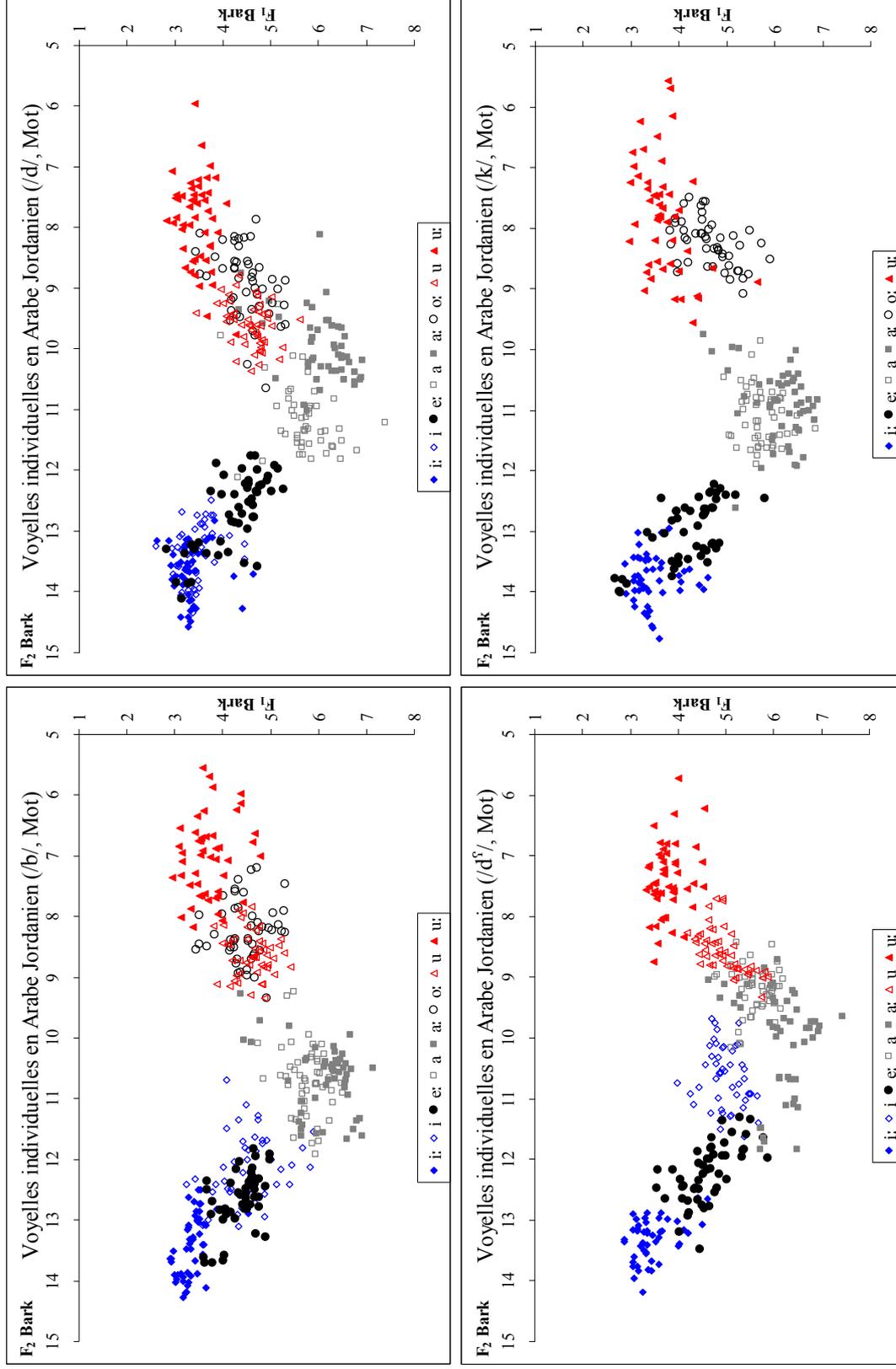


Figure 85 : Dispersion des voyelles en AJ à partir des données individuelles, par lieu d'articulation, en Mot.

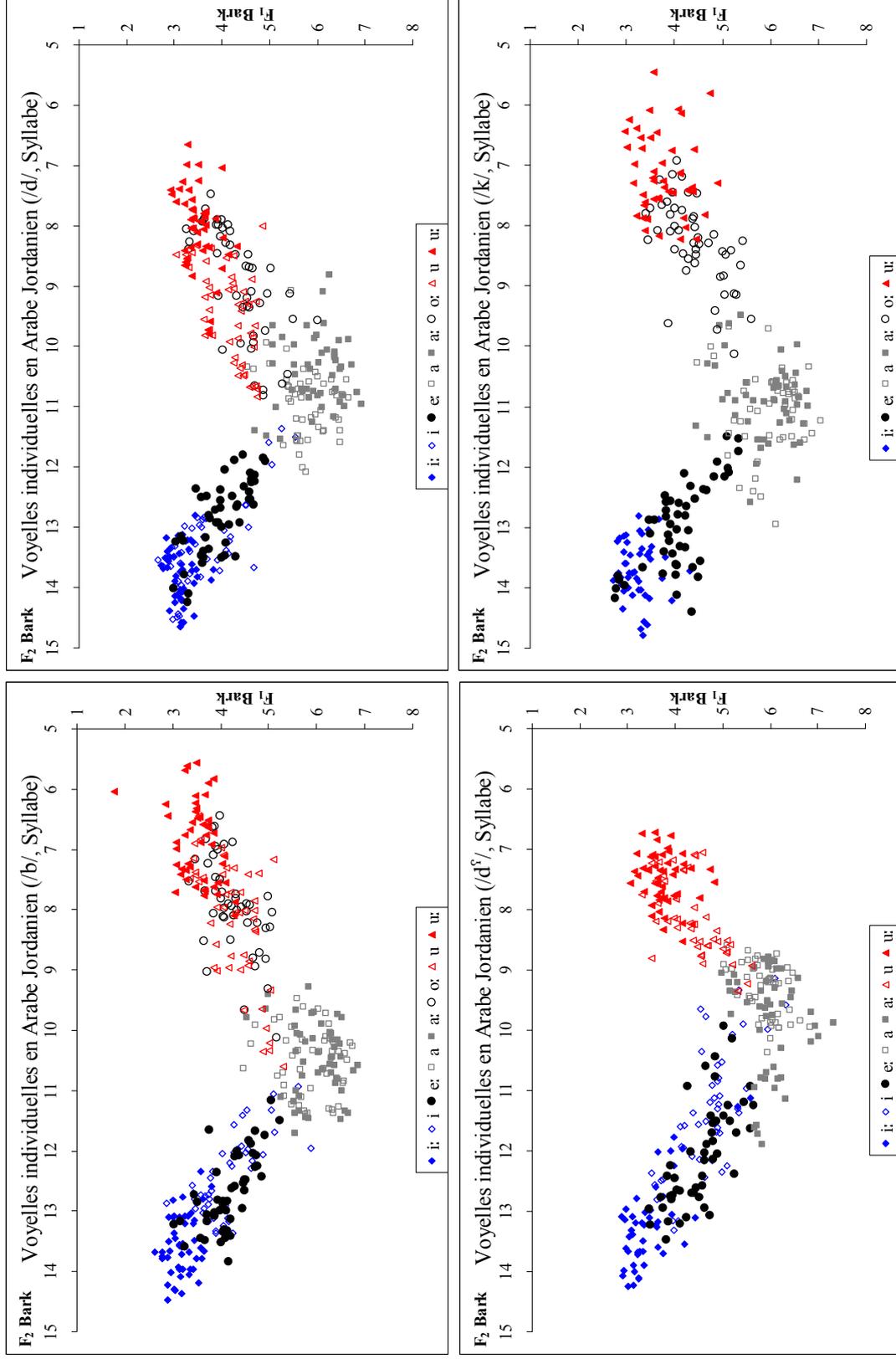


Figure 86 : Dispersion des voyelles en AJ à partir des données individuelles, par lieu d'articulation, en Syllabe.

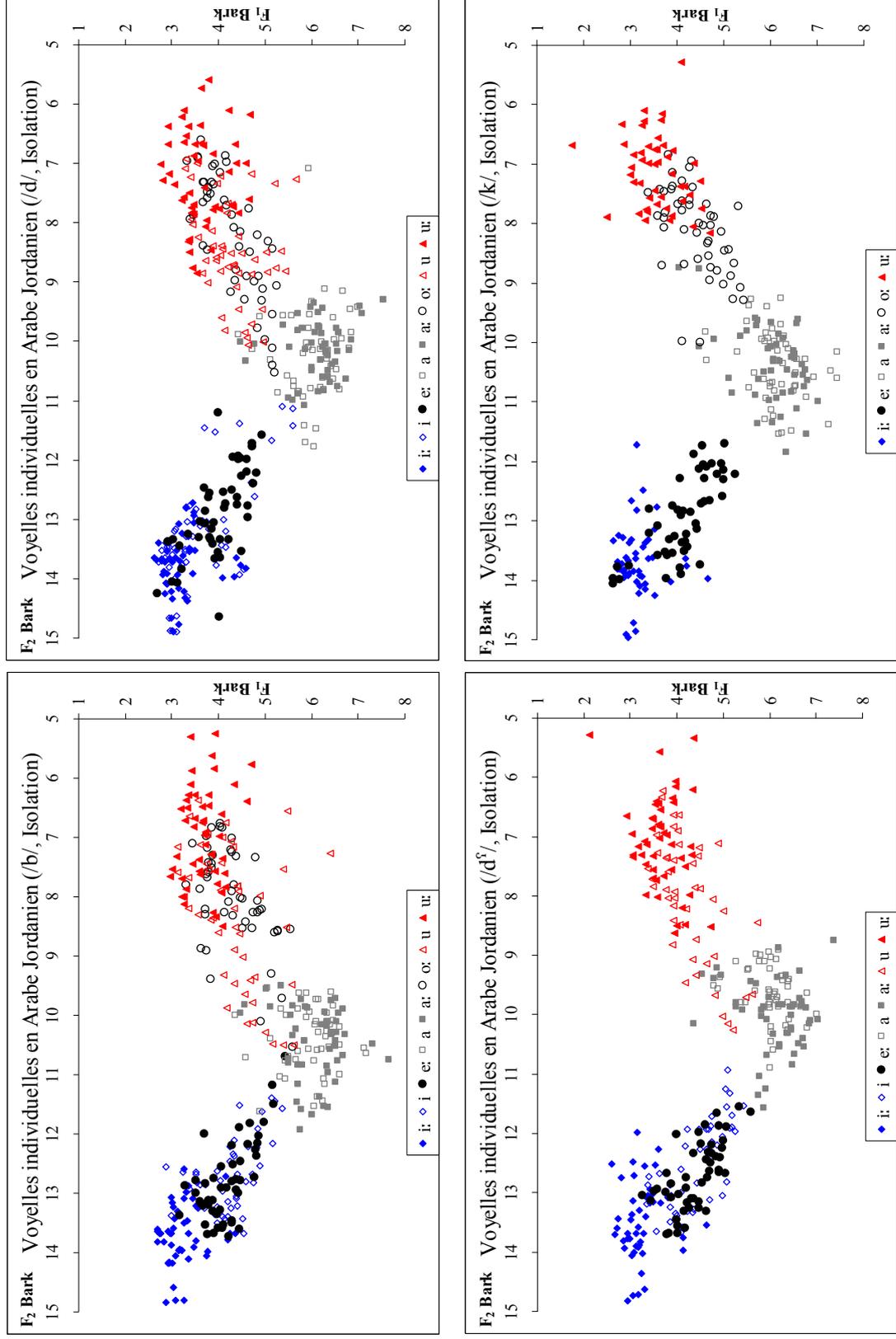


Figure 87 : Dispersion des voyelles en AJ à partir des données individuelles, par lieu d'articulation, en Isolation.

Les représentations graphiques des voyelles à partir des données individuelles montrent l'existence d'une variabilité inter-individuelle importante. Nous nous intéressons à la qualité des voyelles brèves en AJ. Plus particulièrement, la voyelle /i/ dans les trois situations de production et dans les environnements /b d^h/ est réalisée comme une voyelle [e] voire comme un schwa (surtout dans l'environnement /d^h/) ; seule une réalisation de cette voyelle en tant que [i] est observée dans l'environnement /d/. Ceci peut être expliqué par l'item choisi : dans l'environnement /d/, l'item utilisé est [dijja]. En contact avec la semi-voyelle /j/, qui a des fréquences formantiques proches de /i/, les fréquences de la voyelle [e] se voient modifiées pour se rapprocher de celle d'un [i]. La réalisation de la voyelle /u/ brève en AJ se rapproche de celle d'une voyelle [o], dans tous les contextes de réalisation et dans tous les environnements consonantiques (/b d^h/). Quant à la voyelle /a/, nous observons une réalisation de celle-ci plus « centralisée » surtout sur l'axe F₁ par rapport à sa correspondante longue dans tous les environnements consonantiques et dans le contexte Mot. Lorsque nous comparons sa réalisation dans les contextes Syllabe ou Isolation, nous observons un rapprochement très marqué de sa correspondante longue.

4.4.2.2. Résultats en dynamique

Comme en AM, les résultats des **pentés formantiques** obtenues en AJ dans la situation de production Mot (voir Figures 88 à 90), montrent des différences importantes liées aux points de départ des transitions formantiques, à la cible vocalique dérivée et à leur degré d'inclinaison (pour les résultats graphiques en Syllabe et Isolation, voir Annexe 20). Les résultats des **pentés formantiques** de F₁ en AJ (voir Figure 88) montrent des variations importantes en fonction de l'**onset** vocalique dans les environnements /b d k/ : les valeurs sont situées dans une fourchette de 3 à 4,5/5 Bark ; dans l'environnement /d^h/, la dispersion des points de départ des transitions est plus petite, avec des valeurs allant de 3,8 à 4,7 Bark ($F(3, 1400)=39,67$; $p<0,001$). Cette différence peut être expliquée par le degré d'aperture plus important des voyelles (sauf pour les voyelles /a a:/) produites dans cet environnement, avec un rapprochement des valeurs de l'état stable de F₁ : ce rapprochement des valeurs de l'état stable influence directement celles de l'**onset** vocalique. Les résultats des **pentés**

formantiques obtenus pour les deux voyelles problématiques en AJ /i u/, montrent que la voyelle /i/ ne présente ni le même point de départ ni la même cible vocalique dérivée lorsque nous la comparons avec sa correspondante longue dans l'environnement /b/ ; l'**onset** vocalique de la voyelle /i/ est identique à celui obtenu pour la voyelle /e:/. Dans l'environnement /d/, elle présente le même point de départ, la même cible vocalique et le même degré d'inclinaison de sa correspondante longue (ce résultat est en accord avec celui obtenu avec les indices **statiques** et est explicable par l'item utilisé /dijja/). Dans l'environnement /d^h/, nous obtenons le même point de départ uniquement et une cible vocalique significativement différente lorsque nous la comparons avec sa correspondante longue ($F(6, 532)=4,86$; $p<0,001$). Les **pentés formantiques** obtenues sur F_1 pour la voyelle /i/ sont significativement différentes par rapport à celles de sa correspondante longue ; en revanche nous obtenons approximativement des **pentés** identiques entre les deux voyelles /i/ et /e:/ ($F(3, 532)=8,53$; $p<0,001$). Des différences sont observées entre les lieux d'articulation et plus particulièrement dans l'environnement /d/ où les deux voyelles /i/ et /i:/ obtiennent le même degré d'inclinaison, néanmoins dans les environnements /b d^h/, les résultats montrent que les trois voyelles /i/, /i:/ et /e:/ sont significativement différentes ($F(6, 532)=19,41$; $p<0,001$).

La voyelle /u/ présente le même cas de figure : un **onset** vocalique, une cible vocalique dérivée et un degré d'inclinaison différents de ceux de sa correspondante longue dans les deux environnements /b d/ ; dans l'environnement /d^h/, nous observons uniquement le même point de départ ($F(6, 481)=3,06$; $p<0,01$). À partir des résultats des **pentés formantiques** obtenues sur F_1 nous observons des différences importantes entre les deux voyelles brèves et leurs correspondantes longues ; elles se rapprocheraient davantage des voyelles longues /e: o:/.

Les **pentés formantiques** obtenues sur F_2 (voir Figure 89) sont significativement différentes entre les voyelles en AJ et plus particulièrement pour les voyelles problématiques. Ainsi, dans l'environnement /b/, les deux voyelles /i u/ présentent approximativement les mêmes points de départ, les mêmes cibles vocaliques dérivées et des degrés d'inclinaison beaucoup plus forts que les voyelles longues /e: o:/ ($p<0,001$). Dans l'environnement /d/, seule la voyelle /i/ présente approximativement les mêmes points de départ, de cibles

vocaliques et de degré d'inclinaison que sa correspondante longue ; la voyelle /u/ quant à elle se rapproche de la voyelle /o:/ simplement en fonction de la cible vocalique dérivée : les points de départs et les degrés d'inclinaison sont différents de ceux de sa correspondante longue ou de la voyelle /o:/ ($p < 0,001$). Pour l'environnement /d^h/, la voyelle /i/ présente les mêmes points de départ et le même degré d'inclinaison des pentes que la voyelle /o:/ ; la cible vocalique dérivée est différente des deux voyelles /i: o:/ ($p < 0,001$). La voyelle /u/ est produite comme une voyelle /u:/ uniquement à partir des degrés d'inclinaison des **pent**es **formantiques** moyennes ; l'**onset** vocalique et la cible vocalique sont significativement différentes de la voyelle /u:/.

Quant aux résultats sur F₃ (voir Figure 90), les **pent**es **formantiques** obtenues sont significativement différentes plus particulièrement dans l'environnement /b d/, dans l'environnement /d^h/ toutes les pentes de F₃ convergent vers une seule et unique valeur autour des 15 Bark ($F(21, 1400) = 10,62$; $p < 0,001$). Ainsi, les voyelles problématiques en AJ sont différentes de leurs correspondantes longues dans les environnements /b d/ uniquement.

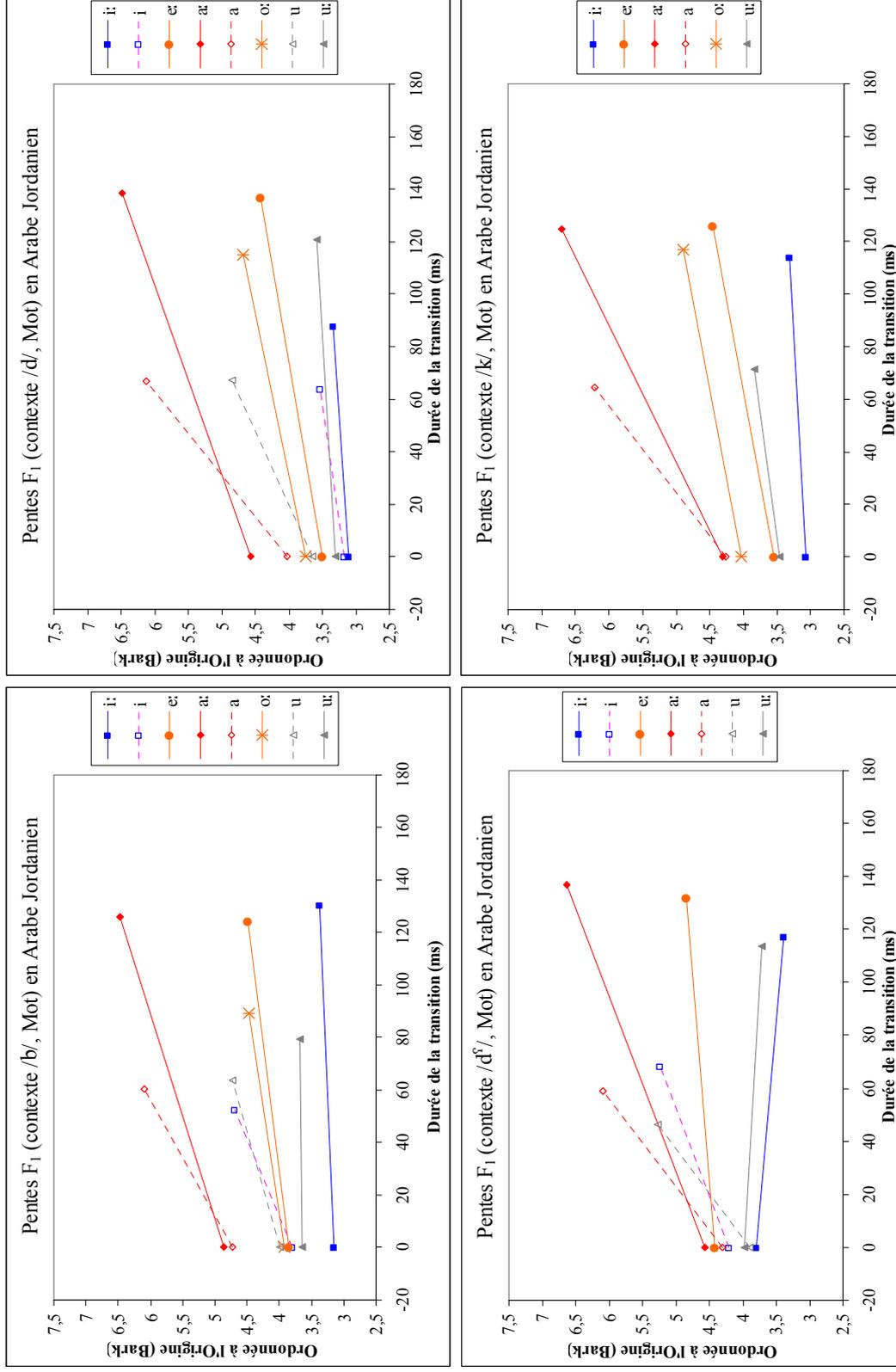


Figure 88 : Pentas formantiques de F₁ des voyelles en AJ sans normalisation du temps, en fonction du lieu d'articulation en Mot.

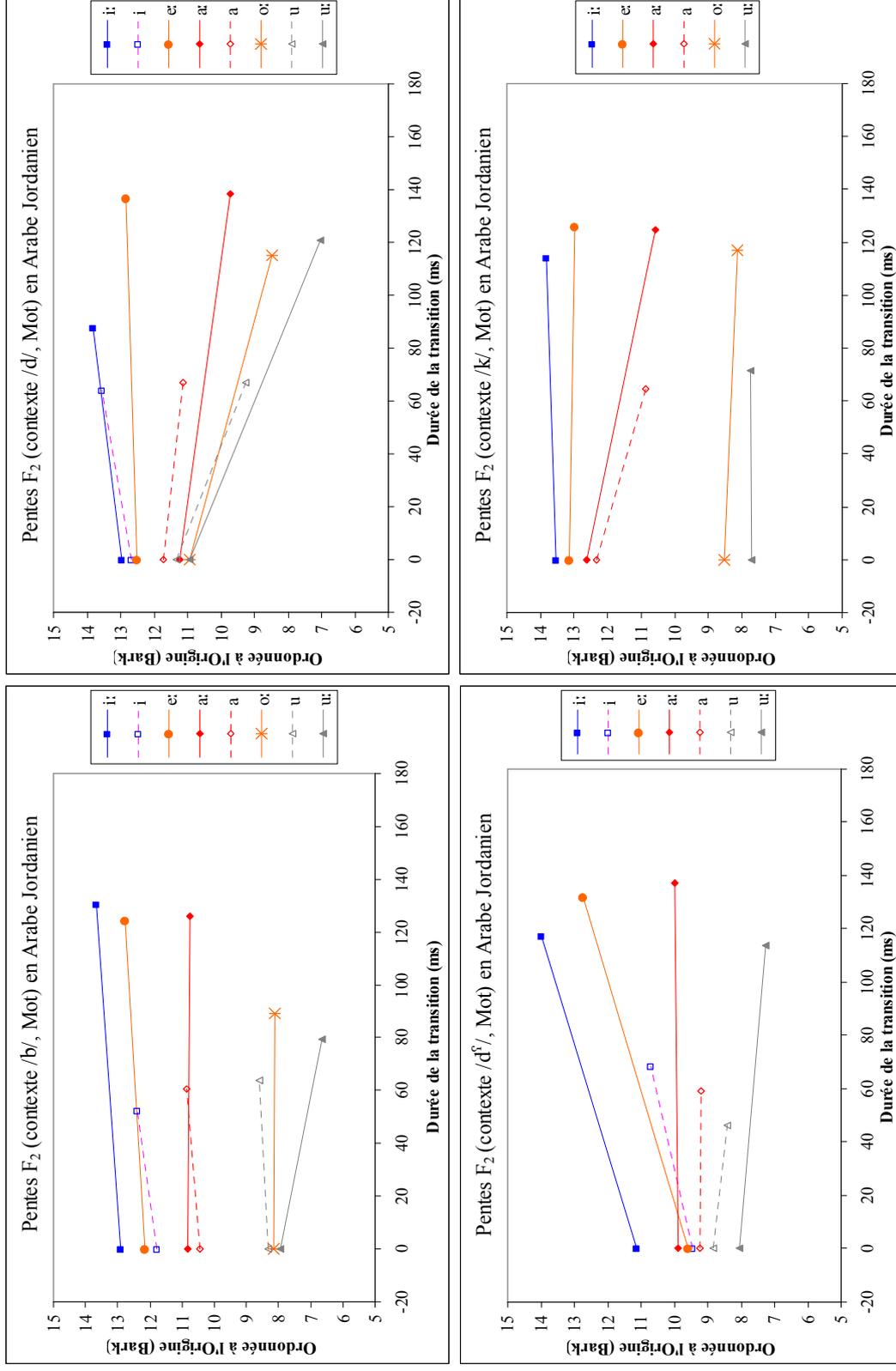


Figure 89 : Pentas formantiques de F₂ des voyelles en AJ sans normalisation du temps, en fonction du lieu d'articulation en Mot.

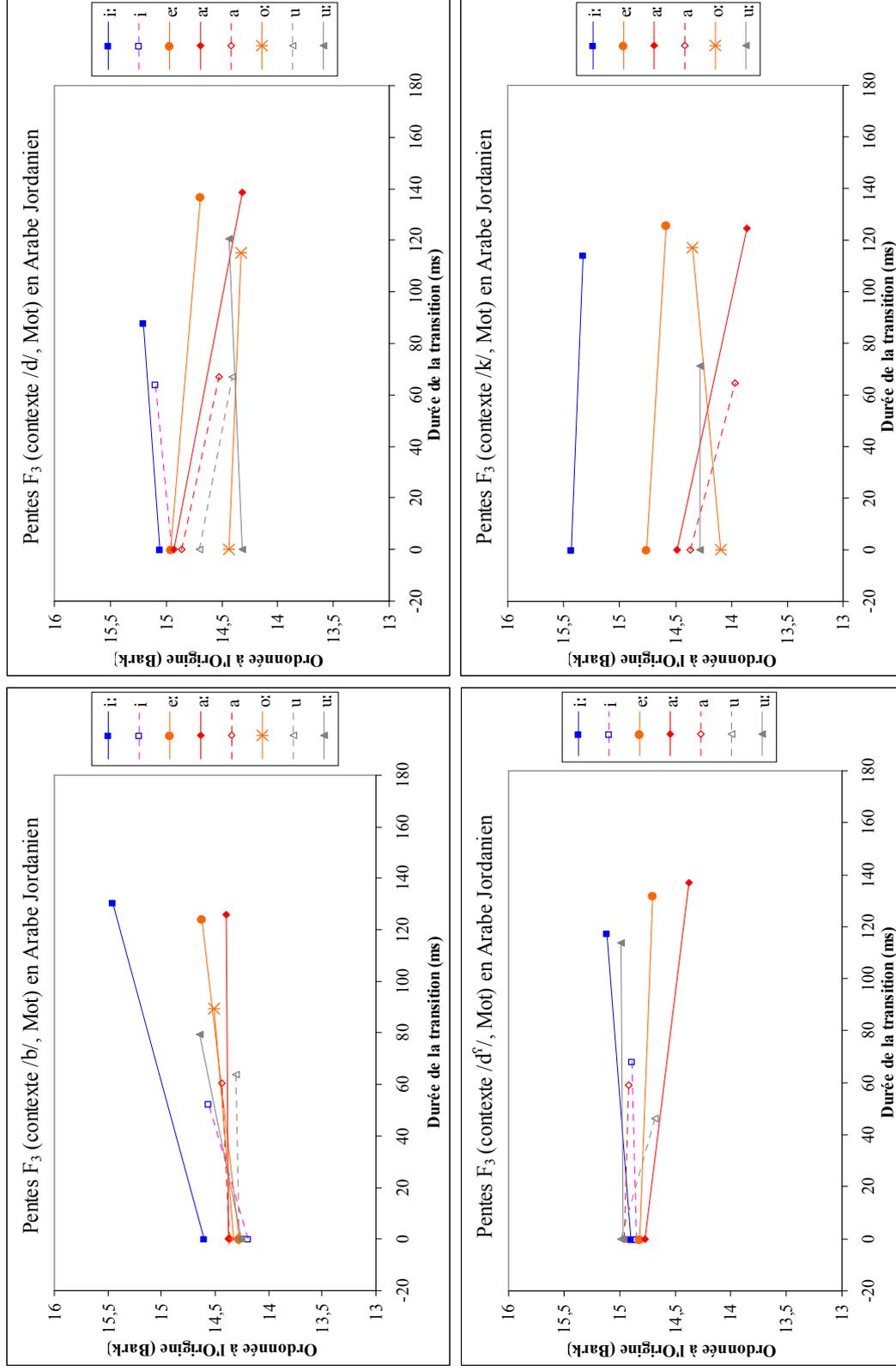


Figure 90 : Pentes formantiques de F₃ des voyelles en AJ sans normalisation du temps, en fonction du lieu d'articulation en Mot.

Les résultats des **pentés formantiques** obtenues en AJ montrent des différences importantes au niveau de la qualité des voyelles brèves /i u/ : est-ce qu'elles se rapprochent qualitativement de leurs correspondantes longues ou est-ce qu'elles sont plus proches des deux voyelles périphériques longues /e: o:/ ? La plupart du temps, nous obtenons un rapprochement entre les deux voyelles /i u/ et les deux voyelles longues /e: o:/.

Globalement, les résultats obtenus avec les deux paramètres **statiques** et **dynamiques** sont homogènes, car nous observons que les deux voyelles brèves /i u/ sont réalisées comme des voyelles mi-fermées, i.e., /e o/ respectivement, en revanche, la voyelle /a/ est réalisée comme un /a/. Nous avons supposé que la structure syllabique des items pouvait affecter le degré d'aperture des voyelles. Ces suppositions ont été validées pour la voyelle /a/ dans la situation Mot, i.e., nous obtenons une voyelle plus « centralisée » avec une réalisation phonétique de type [ɐ] ; dans les deux autres situations de production, cette voyelle est réalisée comme une voyelle [a]. Pour ce qui est des deux voyelles brèves /i u/, leurs réalisations dans les trois situations de production est homogène, car nous obtenons **toujours**, les voyelles /e o/. Autrement dit, la structure syllabique semble n'influencer que la voyelle ouverte. En ce qui concerne les paramètres **dynamiques**, nos résultats ont montré que les deux voyelles brèves /i u/ sont significativement différentes de leurs correspondantes longues en fonction : i) des degrés d'inclinaison des **pentés formantiques**, ii) des points de départ des transitions formantiques et iii) des « cibles vocaliques » dérivées. Nous supposons donc que le système vocalique de l'AJ a bien changé puisque les deux voyelles brèves /i u/ ne sont plus qualitativement identiques à leurs correspondantes longues.

Nous examinerons dans la partie qui suit les résultats de l'analyse discriminante appliquée sur les paramètres **statiques** et **dynamiques** afin d'évaluer d'une part le rôle de ces traits dans la classification des voyelles en AJ et d'autre part, d'examiner les erreurs de classifications qui nous permettront de clarifier cette confusion au niveau de la qualité des voyelles brèves en AJ.

4.4.2.3. Résultats de l'analyse discriminante

4.4.2.3.1. En statique

Nous avons appliqué une analyse discriminante avec validation croisée ayant comme paramètres les fréquences formantiques de F_1 , F_2 & F_3 des voyelles de l'AJ, en incluant et en excluant la durée vocalique. Les taux de classification correcte de cette analyse sont présentés dans le Tableau 23.

Tableau 23 : Taux de classification correcte des voyelles en AJ à partir des valeurs formantiques de F_1 , F_2 & F_3 en fonction de la situation de production et du lieu d'articulation. Les valeurs entre parenthèses indiquent les taux de classification en excluant la durée. (Taux significatifs, Box' M ; $p < 0,001$).

	/b/	/d/	/d ^h /	/k/
Mot	87,3% (68,1%)	85,1% (69,7%)	89,6% (83,2%)	93,2% (78,4%)
Syllabe	66,7% (49,2%)	65,3% (55,4%)	73,9% (60,9%)	72,0% (65,2%)
Isolation	59,4% (48,2%)	56,3% (48,7%)	64,1% (57,1%)	67,7% (66,0%)

Lorsque le paramètre de la durée vocalique est inclus dans l'analyse discriminante, le taux de classification correcte augmente en moyenne de 10% (73,4% avec durée contre 62,5% sans durée), (BinomialQ ; $p < 0,001$). Les taux obtenus dans le Tableau 23 sont assez élevés surtout en Mot ; dans les deux autres situations, les taux sont relativement bas. Cette baisse peut s'expliquer facilement par les confusions dans la classification. Nous présentons sur les Figures 91 et 92 ces confusions en situation Mot (voir Annexes 21 et 22, pour les deux autres situations).

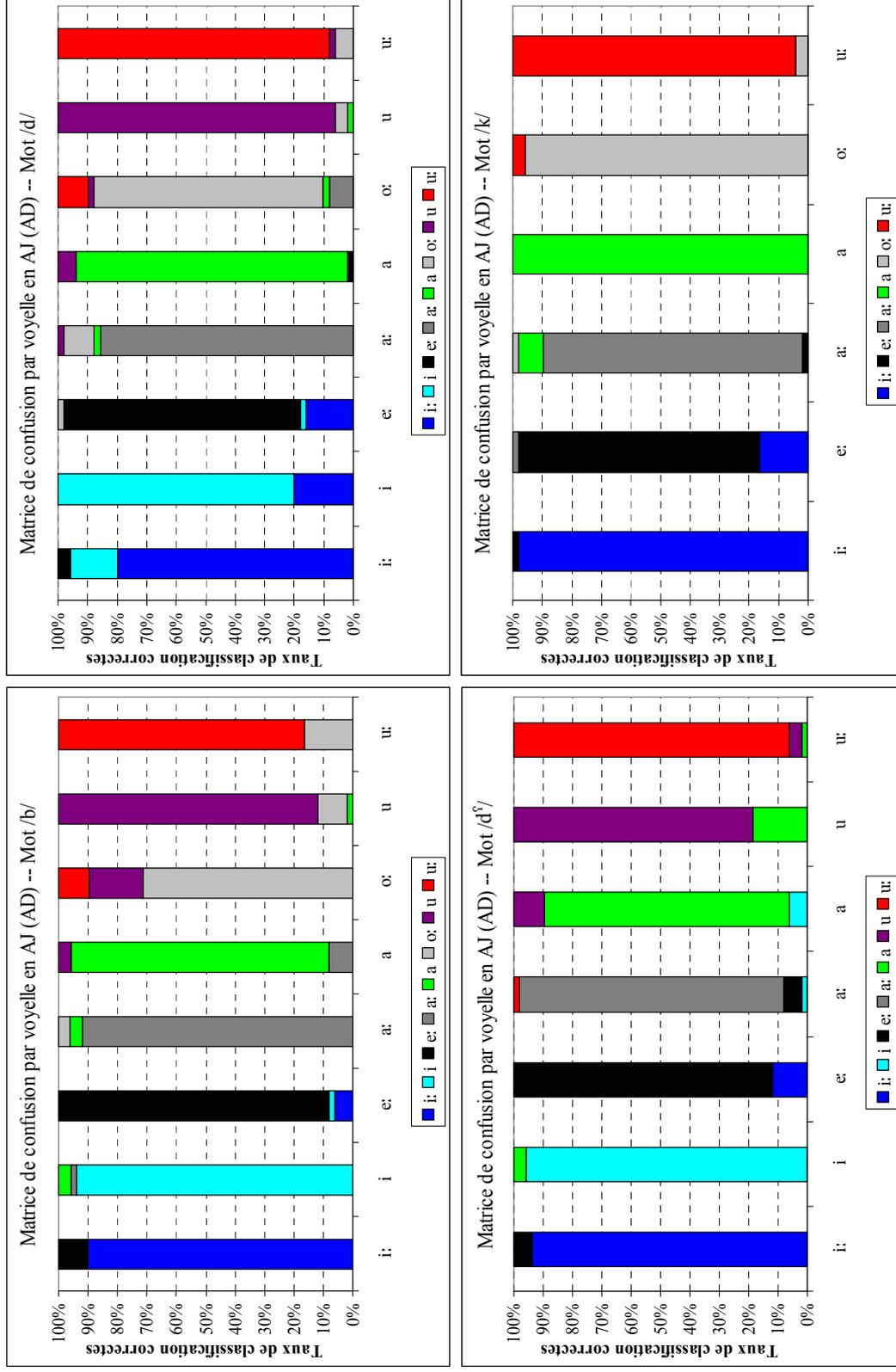


Figure 91 : Matrices de confusion de classification des voyelles de l'AJ en **statique** avec la durée vocalique (AD) en situation de production Mot et en fonction du lieu d'articulation.

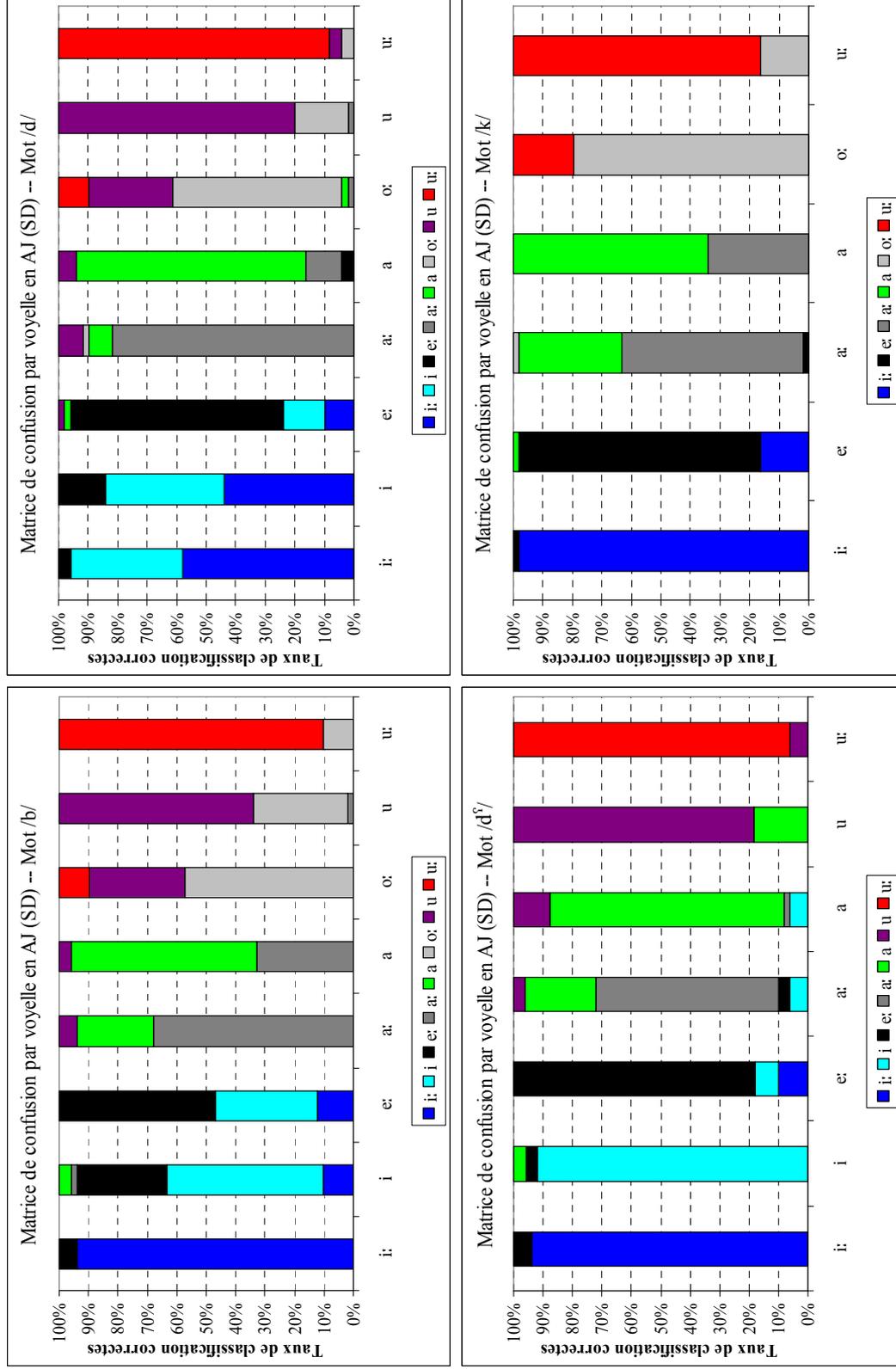


Figure 92 : Matrices de confusion de classification des voyelles de l'AJ en statique sans la durée vocalique (SD) en situation de production Mot et en fonction du lieu d'articulation.

La plupart des confusions obtenues dans les résultats de l'analyse discriminante peuvent s'expliquer par la proximité des voyelles brèves /i a u/ des trois voyelles longues suivantes /e: a: o:/, car elles sont souvent confondues. La voyelle /i/ par exemple n'est jamais confondue avec sa correspondante longue (lorsque le paramètre de la durée vocalique est inclus) sauf dans l'environnement /d/ (à cause de l'entourage contextuel), voir Figure 91. Les deux voyelles longues /i: u:/ obtiennent, en moyenne, plus de 90,0% de classification correcte. Nous obtenons approximativement les mêmes confusions dans les deux autres situations de production : i.e., entre les voyelles brèves /i u/ et les correspondantes longues périphériques /e: o:/ d'un côté ; et entre les deux voyelles /a a:/ de l'autre. Les deux voyelles longues /i: u:/ sont souvent très bien classifiées avec des taux dépassant les 70%.

Nous nous intéressons dans la partie qui suit aux différentes analyses discriminantes appliquées sur les **pentés formantiques** obtenues à partir des différentes analyses de régression. Nous vérifierons ainsi d'une part, la validité de nos résultats et de l'autre l'apport des différentes analyses **dynamiques** dans la description des voyelles de l'AJ.

4.4.2.3.2. En dynamique

4.4.2.3.2.1. Régression linéaire

Les paramètres utilisés dans cette analyse sont les valeurs des **pentés formantiques** et des ordonnées à l'origine de F_1 , F_2 & F_3 des voyelles de l'AJ, ainsi que la durée de la transition formantique. Les résultats de la classification des voyelles de l'AJ en fonction de chaque situation de production et chaque lieu d'articulation sont présentés dans le Tableau 24.

Tableau 24 : Taux de classification correcte des voyelles en AJ à partir de la régression linéaire en fonction la situation de production et du lieu d'articulation. (Taux significatifs, Box's M ; $p < 0,001$).

	/b/	/d/	/d ^s /	/k/
Mot	87,1%	86,1%	89,0%	92,2%
Syllabe	71,2%	66,3%	73,3%	68,2%
Isolation	61,9%	58,8%	68,5%	69,2%

Les taux de classification obtenus sont relativement élevés, toujours dans la situation Mot ; en ce qui concerne les deux autres situations, les taux sont bas. Ceci indique que les trajectoires formantiques dans la situation Mot sont bien modélisées avec l'analyse de régression linéaire. La baisse des taux de classification dans les situations Syllabe et Isolation indique que l'analyse de régression linéaire ne prend pas en compte les trajectoires **intrinsèques** des voyelles.

Les confusions dans la classification des voyelles produites dans la situation Mot, que nous présentons sur la Figure 93 (pour les résultats graphiques des situations de production Syllabe et Isolation, voir Annexe 23), montrent que ce sont les voyelles /i a u/ qui sont toujours confondues avec les voyelles /e: a: o:/ ; les deux voyelles brèves /i u/ ne sont que très rarement confondues avec leurs correspondantes longues et ce dans les trois situations de production. Quant à la voyelle brève /a/, les résultats montrent qu'elle est toujours confondue avec sa correspondante longue et ce quelle que soit la situation de production. Autrement dit, les trajectoires **dynamiques intrinsèque** et **extrinsèque** des deux voyelles /a a:/ sont significativement identiques pour que l'analyse discriminante les confonde.

Nous voulons savoir si la caractérisation des voyelles par Régression non-linéaire, i.e., polynomiales (de 2^{ème} et 3^{ème} ordre) permet de mieux les discriminer en AJ. C'est le sujet de la partie qui suit.

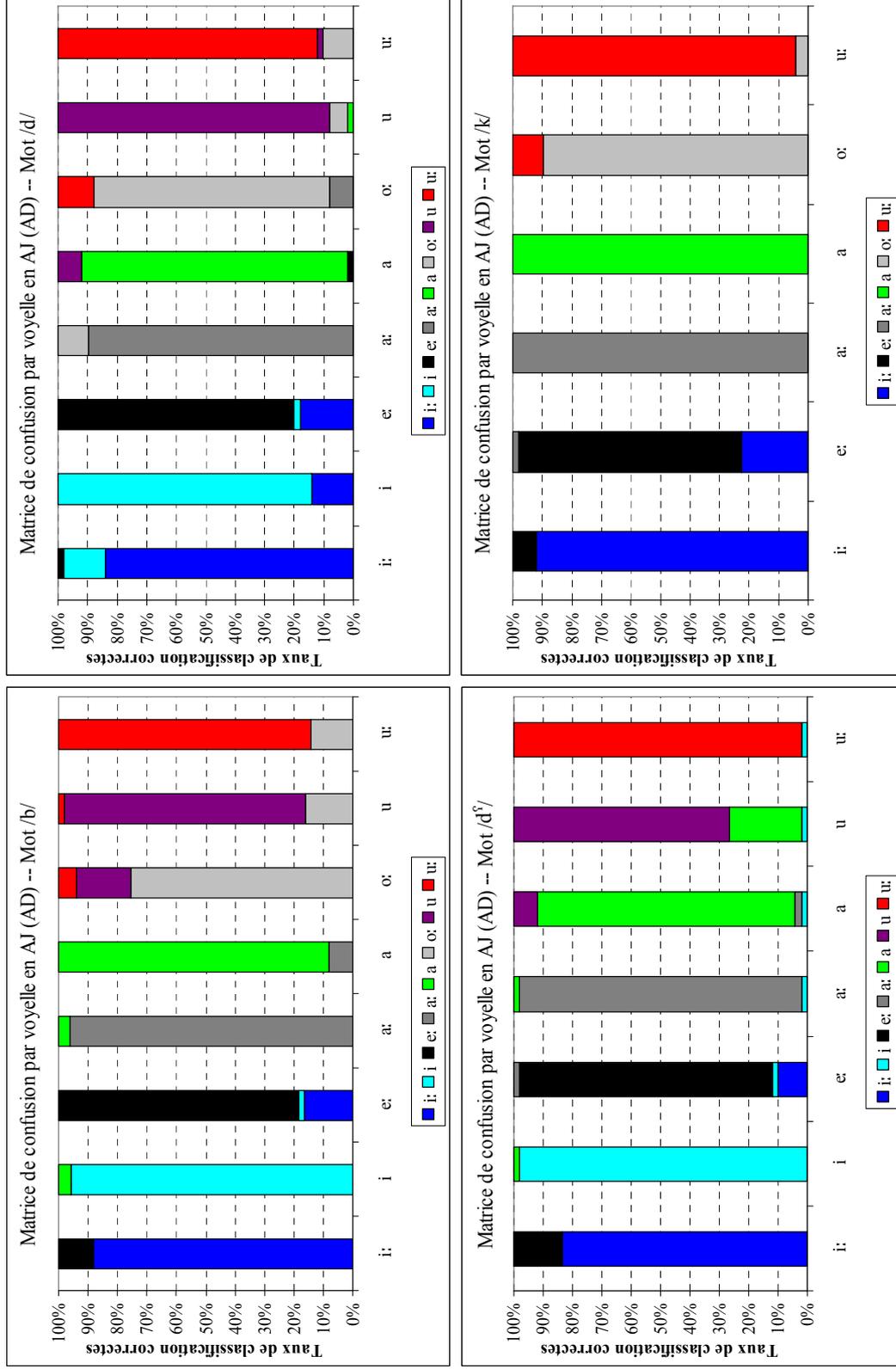


Figure 93 : Matrices de confusion de classification des voyelles de l'AJ en dynamique (régression linéaire) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

4.4.2.3.2.2. Régressions polynomiales de 2^{ème} et 3^{ème} ordre

Les paramètres utilisés pour cette analyse sont : tous les coefficients des régressions polynomiales de 2^{ème} et de 3^{ème} ordre⁶⁴, ainsi que la durée de la transition formantique. Les taux de classification des voyelles de l'AJ en fonction de chaque situation de production et de chaque lieu d'articulation sont présentés dans le Tableau 25.

Tableau 25 : Taux de classification correcte des voyelles de l'AJ à partir des régressions polynomiales de 2^{ème} et de 3^{ème} ordre, en fonction de la situation de production et du lieu d'articulation. Les valeurs entre parenthèses indiquent les taux de classification à partir de l'analyse de régression polynomiale de 3^{ème} ordre. (Taux significatives, Box's M ; $p < 0,001$).

	/b/	/d/	/d ^s /	/k/
Mot	85,8% (83,8%)	83,5% (84,1%)	88,7% (84,9%)	91,9% (90,5%)
Syllabe	69,6% (71,6%)	65,6% (65,8%)	73,0% (70,1%)	66,9% (68,9)
Isolation	64,4% (63,9%)	58,0% (57,0%)	68,1% (65,8%)	68,5% (72,3%)

Ces résultats montrent qu'il est possible de discriminer les voyelles pour un lieu d'articulation et de réalisation donnés avec les coefficients de régressions polynomiales de 2^{ème} et de 3^{ème} ordre, des trois premiers formants ainsi que la durée de la transition formantique. Néanmoins, nous n'observons pas d'amélioration du taux de classification de l'analyse discriminante en augmentant l'ordre de la régression polynomiale. En effet, nous obtenons une baisse, non significative, du taux de classification de 0,5% en moyenne en utilisant un polynôme de 3^{ème} ordre pour toutes les comparaisons (pour la classification des voyelles, le taux est de 73,7% avec la régression polynomiale de 2^{ème} ordre contre 73,2% en 3^{ème} ordre), (BinomialQ ; $p=1$). Il semblerait que l'utilisation de la régression polynomiale de 3^{ème} ordre n'apporte pas d'amélioration à la caractérisation dynamique des voyelles en AJ comme en AM. Il est évident également que l'utilisation de la régression polynomiale dans la situation de production Mot en AJ apporte davantage de précision dans la caractérisation **dynamique** des voyelles (i.e., les effets de coarticulation) par rapport au **statique** : en Syllabe

⁶⁴ Les coefficients utilisés sont : pour la régression polynomiale de 2^{ème} ordre, a_0 = la moyenne des valeurs, a_1 = la droite de régression et a_2 = l'ordonnée à l'origine ; pour la régression polynomiale de 3^{ème} ordre, a_0 = la moyenne des valeurs, a_1 = la droite de régression, a_2 = l'excursion parabolique de la voyelle indépendamment de sa droite et a_3 = l'ordonnée à l'origine.

et Isolation, les taux de classification sont largement inférieurs. Ces constats n'ont pas été confirmés en AM.

Comme avec la régression linéaire, la plupart des confusions dans les résultats de l'analyse discriminante sont dues aux différences de qualités des voyelles brèves en AJ. Pour les deux analyses de régressions polynomiales de 2^{ème} et de 3^{ème} ordre, nous obtenons les mêmes tendances ; i.e., des confusions dues au rapprochement entre les deux voyelles brèves /i u/ des voyelles périphériques longues /e: o:/. Afin d'examiner ces confusions, nous présentons sur les Figures 94 et 95 ces taux en situation de production Mot (pour les résultats graphiques des confusions en Syllabe et en Isolation, voir Annexes 24 et 25). Nous obtenons une homogénéité dans les confusions de classification des voyelles de l'AJ avec les deux situations de production, i.e., ce sont les mêmes voyelles qui sont toujours confondues entre elles. Nous observons que les taux de confusions augmentent avec la régression polynomiale de 3^{ème} ordre, ce qui explique la baisse des taux de classification des voyelles de l'AJ.

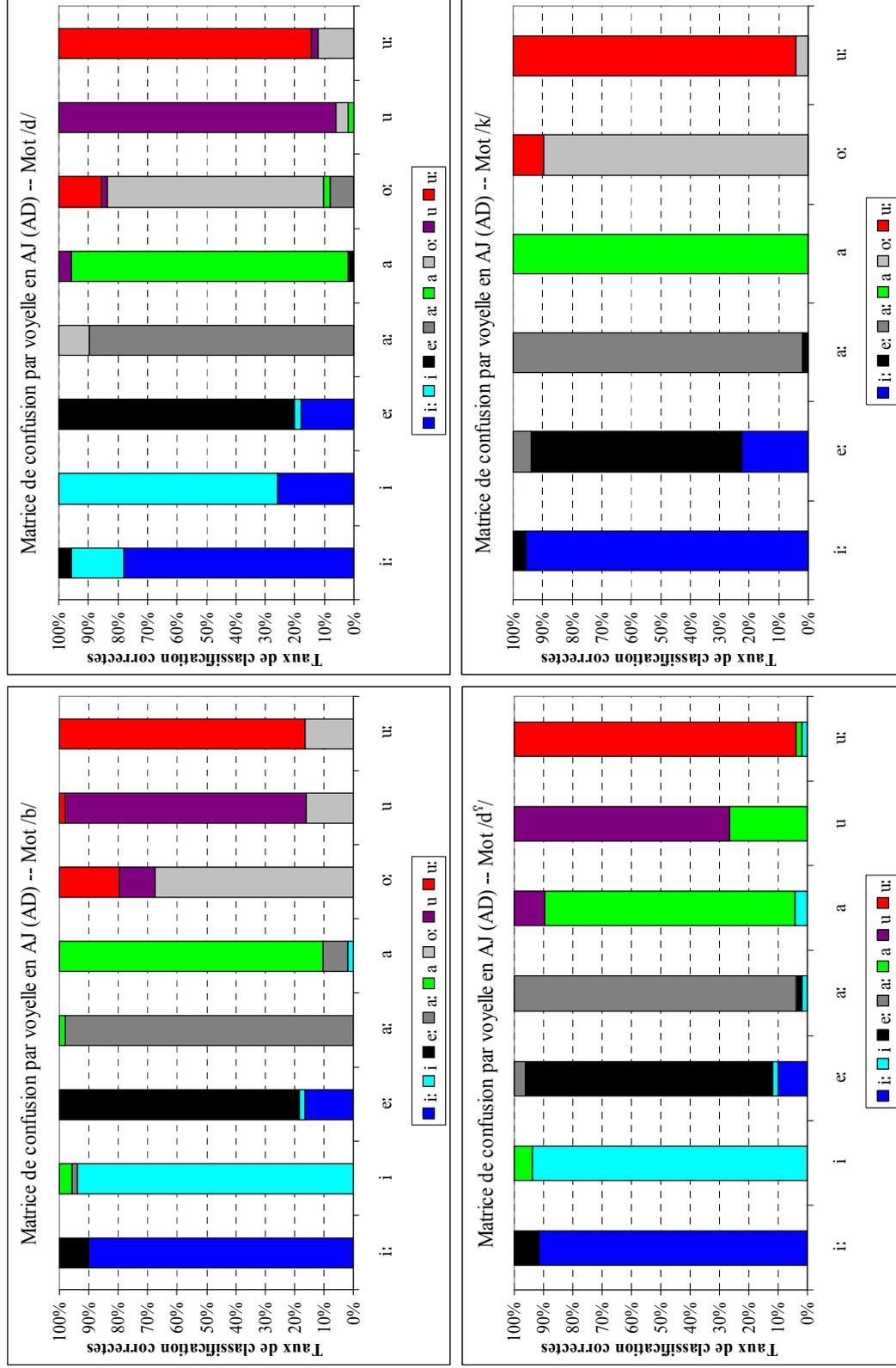


Figure 94 : Matrices de confusion de classification des voyelles de l'AJ en dynamique (régression polynomiale de 2^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

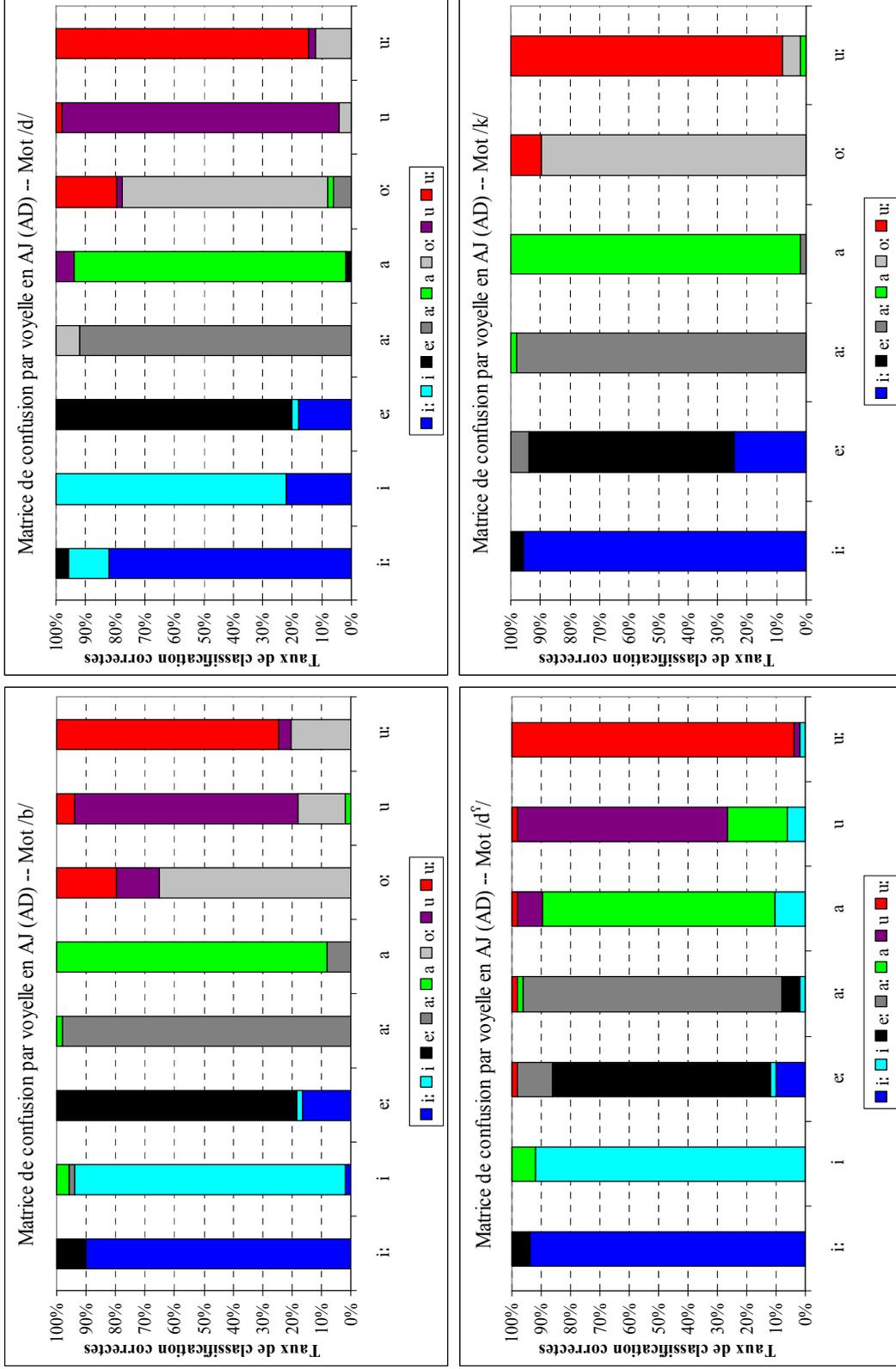


Figure 95 : Matrices de confusion de classification des voyelles de l'AJ en dynamique (régression polynomiale de 3^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

4.4.2.4. Discussion des résultats en statique vs en dynamique

Les résultats obtenus en AJ à partir des deux analyses **statique** et **dynamique** montrent que les deux voyelles brèves /i u/ sont positionnées dans l'espace vocalique à proximité directe des deux voyelles longues /e: o:/, respectivement et pas du tout aux côtés de leurs correspondantes longues. Les **pentés formantiques** obtenues pour les deux voyelles brèves /i u/ montrent leurs différences avec leurs correspondantes longues en fonction du degré d'inclinaison des transitions ; du point de départ des transitions et de la valeur des « cibles vocaliques » dérivées. Leurs **pentés formantiques** sont proches de celles des deux voyelles longues /e: o:/. Ainsi, du point de vue **statique** et **dynamique**, les deux voyelles brèves /i u/ ont changé de degré d'aperture en AJ et par conséquent elles sont produites par les locuteurs jordaniens comme /e o/ (dans le mot [dijja], la voyelle /i/ est produite comme [i], à cause de la semi-consonne [j] qui présentent les mêmes structures acoustiques que la voyelle /i/). La voyelle brève /a/ a été produite par les locuteurs jordaniens comme une voyelle mi-ouverte [ɐ] dans la situation de production Mot et comme une voyelle ouverte [a] dans les situations Syllabe et Isolation, se rapprochant ainsi sa correspondante longue /a:/.

Les différences observées ont permis de mettre en évidence les particularités du système vocalique de l'AJ ; différences qui ont été validées par l'analyse discriminante. En effet, les résultats de cette analyse que nous venons de présenter, ont montré que le rapprochement entre les deux voyelles brèves /i a u/ des deux voyelles longues /e: a: o:/ respectivement était à l'origine de la plupart des confusions dans les analyses. Outre les différences de timbres vocaliques en AJ, l'utilisation des différentes analyses **dynamiques** a permis de quantifier les trajectoires formantiques **intrinsèques** et **extrinsèques**. La comparaison entre les taux moyens de la classification des voyelles de l'AJ présentée sur la Figure 96, permet de mesurer les effets de la **dynamique**.

La comparaison entre les taux obtenus à partir des « cibles vocaliques » avec et sans la durée vocalique montre que lorsque la durée vocalique est ajoutée à l'analyse discriminante, le taux moyen par situation de production augmente significativement de 14% en situation

Mot, de 9% en Syllabe et de 7% en Isolation, (BinomialQ ; $p < 0,001$). La comparaison entre les taux obtenus à partir des « cibles vocaliques » sans la durée vocalique et les trois analyses **dynamiques** montre que les taux de ces dernières augmentent significativement en **dynamique** de 12% lorsque les coefficients de la régression linéaire sont utilisés (BinomialQ ; $p < 0,001$) et de 11% lorsque chacune des analyses de régressions polynomiales est utilisée (BinomialQ ; $p < 0,001$). Ces résultats vont dans le sens d'une meilleure caractérisation des voyelles de l'AJ en **dynamique** : caractérisation mettant en évidence les trajectoires formantiques.

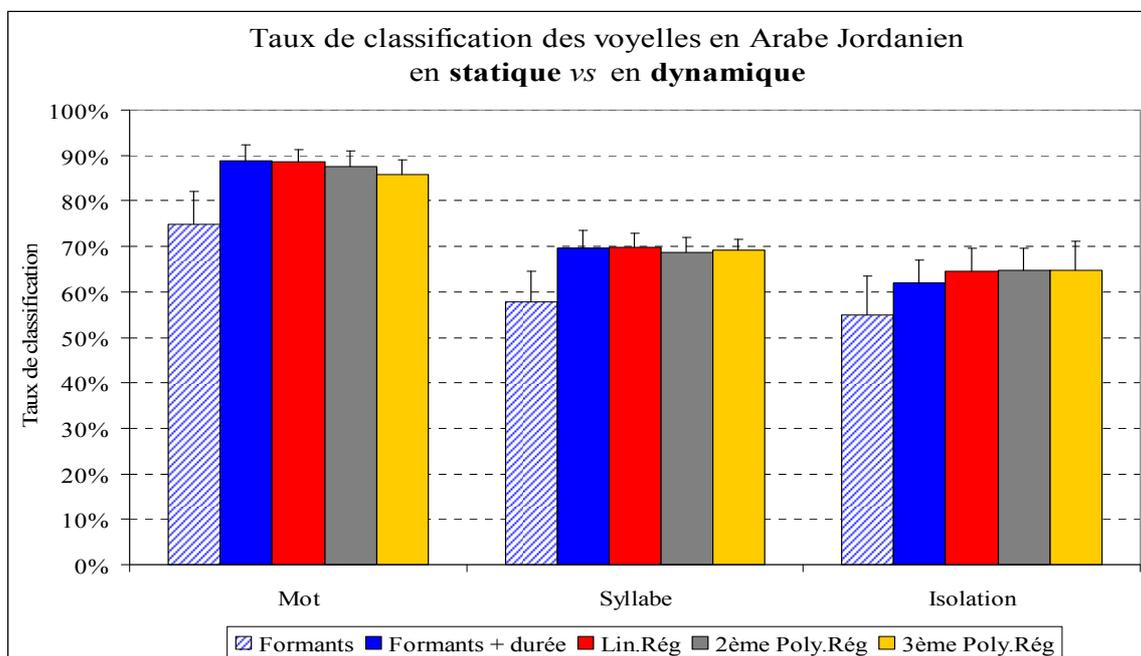


Figure 96 : Taux de classification correcte moyens des voyelles de l'AJ produites dans les contextes /b d d^f k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur).
 Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} ou 3^{ème} ordre).

Toutefois, nous observons que le taux moyen de classification des voyelles de l'AJ, obtenu en utilisant les valeurs des « cibles vocaliques » et la durée, en situation de production Mot est significativement supérieur de 2% à celui obtenu à partir des deux analyses de régressions polynomiales (BinomialQ, $p < 0,001$) ; aucune différence n'est obtenue en comparaison avec la régression linéaire. En situation Syllabe, les taux obtenus à partir des trois analyses de régression en comparaison avec ceux obtenus à partir des « cibles vocaliques » avec la durée ne sont pas significativement différents. Enfin, en situation Isolation, nous obtenons une augmentation significative du taux de classification moyen

quelle que soit l'analyse de régression utilisée de 3% en comparaison avec les taux obtenus à partir de la classification des voyelles en utilisant les « cibles vocaliques » et la durée (BinomialQ, $p < 0,001$).

Ces résultats indiquent que le paramètre de la durée vocalique semble être un indice primordial dans la caractérisation des voyelles de l'AJ. En effet, nous avons expliqué que lorsque la durée vocalique est ajoutée à l'analyse discriminante, elle a été toujours utilisée en premier lieu comme facteur discriminatoire entre les voyelles en AJ. L'utilisation de la **dynamique** donne une bonne caractérisation des trajectoires **dynamiques intrinsèques**, car elle permet de mieux modéliser les trajectoires formantiques propres à la voyelle et plus particulièrement dans la situation de production Isolation. Nous pensons donc que ces paramètres dynamiques permettront aux auditeurs jordaniens de mieux discriminer les voyelles de leur système (voir Chapitre 5).

Les résultats obtenus dans les parties précédentes nous conduisent à proposer le système vocalique en AJ comme ayant les huit voyelles suivantes : /i: e e: a a: o: o u:/. Ce système prend en compte les variations dues aux différents environnements consonantiques et structures syllabiques.

4.4.3. *En Français*

Dans cette partie, nous décrivons le système vocalique du FR en commençant par présenter les analyses entreprises en **statique** et ensuite en **dynamique**, en nous concentrant sur les différences de « cibles vocaliques » et de **pentés formantiques** observés en fonction des situations de production et des lieux d'articulation des consonnes adjacentes. Nous nous intéresserons plus particulièrement au statut des voyelles problématiques du FR, à savoir les voyelles dites « médianes » : /e ε ø œ o ɔ/ et des deux voyelles ouvertes /a α/.

Nous nous intéressons à ces voyelles, car elles connaissent de grandes variations dialectales et régionales dans leurs productions et leur statut phonologique. Dans leur étude sur le statut des deux voyelles ouvertes /a α/ en FR, Arnaud & Konopckzynski (1997*) ont montré que les locuteurs âgés conservent encore la distinction entre les deux voyelles du

français /a/ et /ɑ/, dans la région Haut-Jurassienne, tandis que les locuteurs jeunes ne conservent qu'un /a/ standard. Les explications avancées par les auteurs trouvent leurs origines en sociolinguistique. En effet, Arnaud & Konopckzynski (1997*) explique qu'étant donné que les déplacements dans les régions affectent généralement les populations jeunes, ce sont donc les locuteurs âgés qui conserveront au mieux les variétés régionales et dialectales. Ces différences semblent être largement présentes dans d'autres villes françaises (Walter, 1982). En effet, Walter (1982) a proposé des analyses phonologiques des systèmes vocaliques et consonantiques des idiolectes français par 111 informateurs regroupés en 35 régions. Les principaux résultats qu'elle a obtenus confirment l'existence d'une division géographique en différentes zones dialectales. Une des variations existantes concerne les voyelles dites médianes. Cette différence est notée dans les régions Nord~Sud en comparaison avec la région Centre. En effet, en français dit standard, 6 phonèmes vocaliques d'aperture moyenne existent : /e ε ø œ o ɔ/. Selon les régions, ces 6 timbres vocaliques vont correspondre soit à 6 phonèmes (avec 6 variations contextuelles), soit à 3 phonèmes (avec 6 réalisations phonétiques). Au niveau phonétique, ces 6 voyelles sont en distribution complémentaire, selon le contexte syllabique. Ainsi, la voyelle mi-fermée est produite en syllabe ouverte et la mi-ouverte en syllabe fermée (i.e., [e o ø] en syllabe ouverte et [ε ɔ œ] en syllabe fermée). Les locuteurs du centre possèdent six phonèmes pour ces six sons phonétiques. Le Tableau 26 montre un résumé des distinctions au niveau phonémique des voyelles médianes entre les deux zones régionales en France : Nord~Sud et Centre.

Tableau 26 : Schématisation des distinctions entre les deux zones régionales en France au niveau de la réalisation des voyelles médianes. (Adapté de Walter, 1982).

	CENTRE : 6 phonèmes	NORD/SUD : 3 phonèmes
[e]/[ε]	CV [e]~[ε] "pré" vs "près" CVC [ε] "prête"	CV [e] "lait", "lé" <i>neutralisation de l'opposition</i> CVC [ε] "laide"
[o]/[ɔ]	CV [o] "pot" CVC [o]~[ɔ] "saule" vs "sol"	CV [o] "pot" <i>neutralisation de l'opposition</i> CVC [ɔ] "sol", "saule"
[ø]/[œ]	CV [ø]~[œ] "feu" vs. "de" CVC [œ] "jeune" <i>sauf contextes spécifiques :</i> [-z] "heureuse" [-t] "émeute", "meute"	CV [ø] "feu", "de" <i>neutralisation de l'opposition</i> CVC [œ] "jeune", "émeute" Sauf à Lyon CVC [ø] "jeune"

Dans notre étude, nous n'allons pas étudier les voyelles « médianes », nous vérifierons néanmoins leurs réalisations par les différents locuteurs. Commençons tout d'abord par les résultats obtenus en **statique**.

4.4.3.1. Résultats en statique

Nous présentons sur la Figure 97 les résultats graphiques de la dispersion des timbres vocaliques moyens en FR en fonction des trois situations de production : Mot, Syllabe et Isolation. Les résultats statistiques ne montrent pas un effet significatif d'un espace vocalique plus grand en Isolation par rapport aux deux situations Mot et Syllabe, en revanche, cet effet est significatif pour l'interaction entre les situations de production et les voyelles : sur F_1 , les voyelles /i y u e o/ sont significativement plus « fermées » et les voyelles /ε ø œ ɔ a α/, significativement plus « ouvertes » en situation Isolation par rapport au Mot ou en Syllabe ($F(20, 4699) = 2,43$; $p < 0,001$) ; sur F_2 , où l'effet est beaucoup plus visible, les voyelles /i e ε/ sont significativement plus « antérieures » et les voyelles /y ø œ u o ɔ a α/, significativement plus « postérieures » en Isolation par rapport aux situations Mot ou Syllabe ($F(20, 4699) = 24,18$; $p < 0,001$).

L'aire moyenne obtenue dans chaque situation de production va dans le sens de l'hypothèse d'un espace vocalique plus grand en situation Isolation par rapport celui en Mot ou en Syllabe : 10,52 Bark² en Mot, 10,81 Bark² en Syllabe et 13,31 Bark² en Isolation (différences significatives entre les situations Isolation et Mot ($F(1, 18) = 5,42$; $p < 0,05$) et une tendance à observer un espace vocalique plus grand en Isolation par rapport au Syllabe ($F(1, 18) = 3,77$; $p = 0,068$) ; aucune différence n'est obtenue entre les deux situations Mot et Syllabe ($F(1, 18) = 0,06$; $p = 0,81$)).

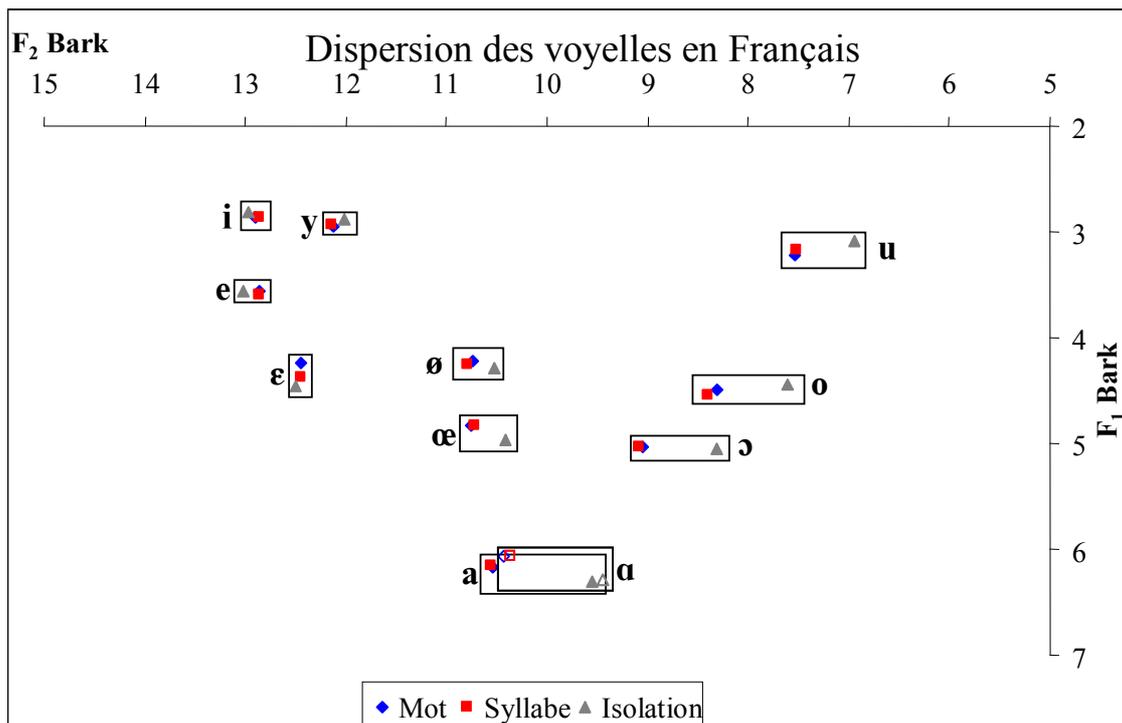


Figure 97 : Espace vocalique en FR en fonction de la situation de production.

En ce qui concerne les résultats de l'interaction entre les situations de production et les lieux d'articulation des consonnes adjacentes ; les résultats graphiques, présentés sur la Figure 98, montrent que les espaces vocaliques obtenus en situation de production Isolation sont significativement les plus grands, i.e., plus « périphériques » pour les trois lieux d'articulation ; ceux obtenus en Mot qui sont significativement les plus petits, i.e., plus « centralisés », uniquement sur F_2 ($F(4, 4699) = 7,42 ; p < 0,001$). Ces résultats restent homogènes pour l'interaction entre les situations de production, les lieux d'articulation et les voyelles, i.e., effet d'agrandissement de l'espace en Isolation par rapport à la situation Mot, sur F_2 uniquement ($F(40, 4699) = 1,86 ; p < 0,001$).

RÉSULTATS EN PRODUCTION DES VOYELLES

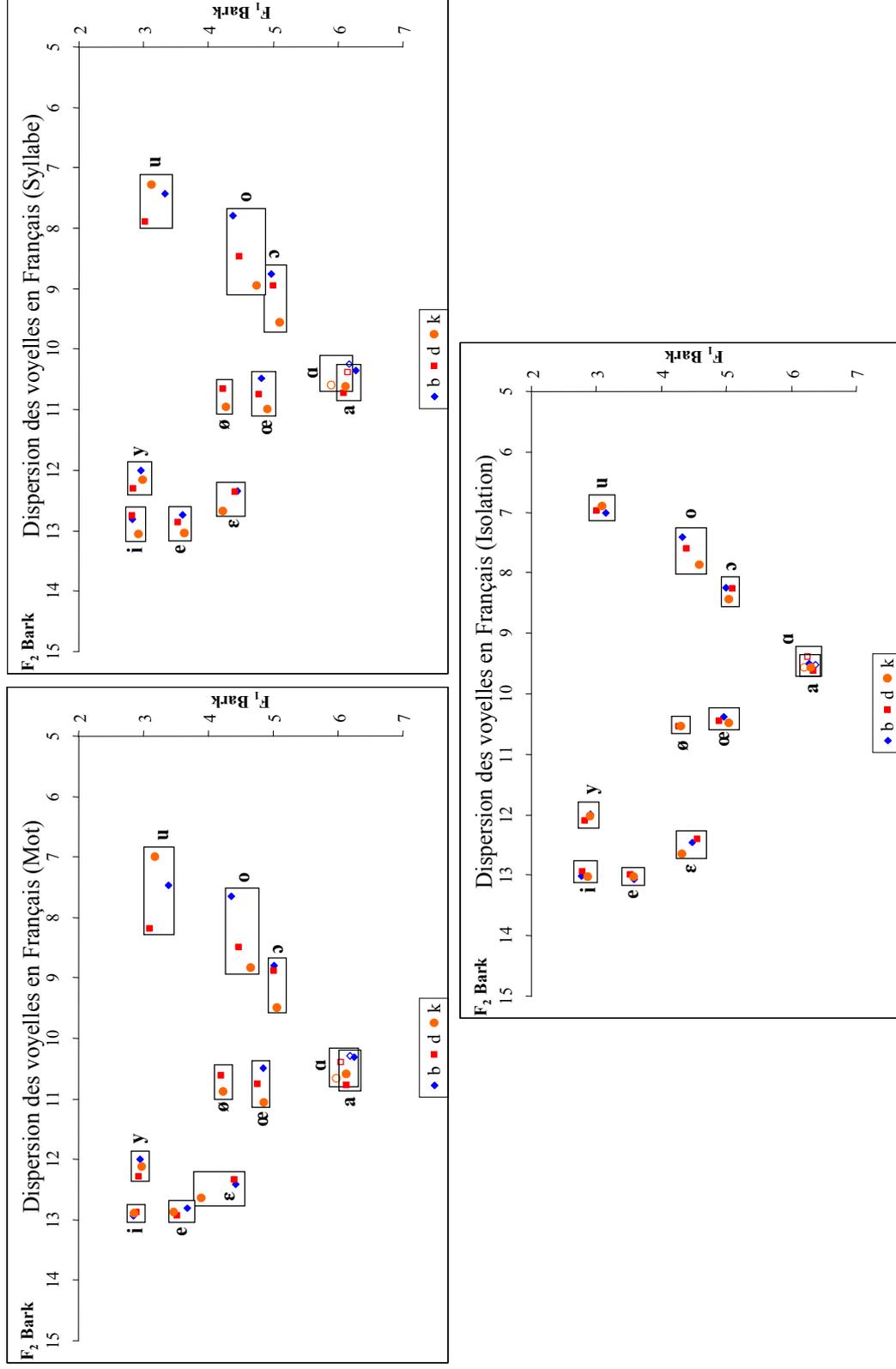


Figure 98 : Espaces vocaliques en FR obtenus en fonction de chaque situation de production et du lieu d'articulation des consonnes /b d k/.

Les aires de dispersion obtenues en FR confirment l'hypothèse selon laquelle l'espace vocalique en situation Isolation est plus grand que celui en Mot ($F(2, 81)=7,49$; $p<0,005$) ; et aucune différence n'est obtenue entre les aires de dispersion en Mot et en Syllabe. Nous nous attendions à n'obtenir pas de différence entre ces deux situations de production, car nous avons observé que les locuteurs français avaient tendance à produire les voyelles dans ces deux situations de production de la même façon. En ce qui concerne l'interaction entre l'aire et le lieu d'articulation, les résultats statistiques ne donnent pas de différences significatives, i.e., quel que soit le lieu d'articulation, en situation Isolation, l'espace vocalique est plus grand que celui en Mot ou en Syllabe.

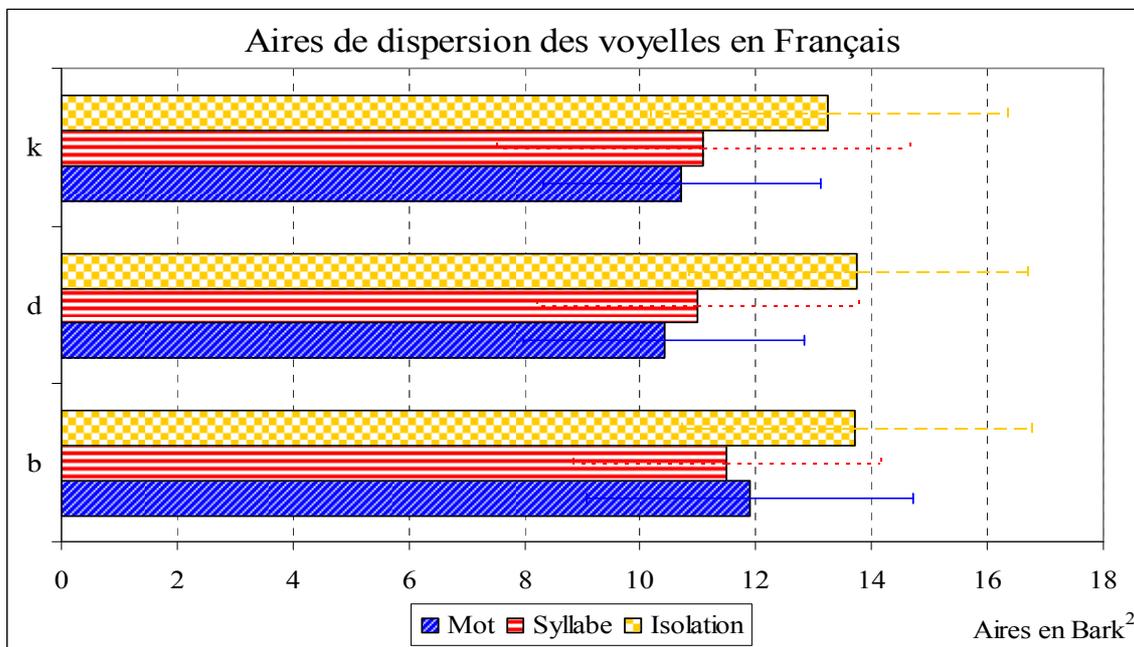


Figure 99 : Aires de dispersion (moyenne et écart-type) des voyelles du FR en fonction de la situation de production et du lieu d'articulation.

Intéressons-nous maintenant aux voyelles problématiques du FR. Les résultats obtenus montrent de faibles différences significatives entre les deux voyelles ouvertes /a/ et /ɑ/ en FR : sur F_1 , la voyelle /ɑ/ est légèrement plus « fermée » que /a/ ($F(1, 882)=4,09$; $p<0,05$) ; et sur F_2 , la voyelle /ɑ/ est légèrement plus « postérieure » que /a/ ($F(1, 882)=8,33$; $p<0,005$). L'interaction avec les trois situations de production s'est révélée non significative sur les deux axes, indiquant l'homogénéité des différences entre les deux voyelles (sur F_1 , ($F(2, 882)=0,29$; $p=0,75$) ; et sur F_2 , ($F(2, 882)=0,52$; $p=0,59$)). L'environnement /k/ antérieure

les deux voyelles ouvertes tandis que dans l'environnement /b/, ces deux voyelles sont postériorisées, ($F(2, 882) = 7,54$; $p < 0,001$), et quelle que soit la situation de production. Cet effet n'est pas attendu, car en générale, c'est l'environnement /k/ qui semble postérioriser l'espace vocalique plus que l'environnement /b/. La position des trois paires de voyelles intermédiaires /e~ɛ/, /ø~œ/ et /o~ɔ/ est différente en fonction du lieu d'articulation ($p < 0,001$) et de la situation de production ($p < 0,001$), néanmoins, ces voyelles restent distinctes au niveau des productions des locuteurs⁶⁵.

Nous avons calculé la taille des ellipses de dispersion de chaque voyelle en fonction de la situation de production et du lieu d'articulation. Nous observons des ellipses de dispersion (voir Figure 100 et Tableau 27) plus grandes pour les voyelles /ɛ a ɔ o ø œ/ dans les trois situations de production. Les voyelles produites dans la situation de production Isolation obtiennent les ellipses les plus petites comparées aux mêmes voyelles produites dans les situations Mot et/ou Syllabe (sauf pour /ɛ/ où l'on obtient l'effet inverse). Ces résultats vont dans le sens de l'hypothèse selon laquelle les voyelles produites en Isolation obtiennent des réalisations homogènes entre les locuteurs avec peu de chevauchements dans leurs réalisations. Ce qui conduit à obtenir des ellipses de dispersion intra-catégorie vocalique petites. Les trois voyelles /i e y/ obtiennent les ellipses de dispersion les plus petites dans les trois situations de production ; les chevauchements maximaux sont obtenus pour les deux voyelles /o ɔ/, qui sont dûs aux différences de réalisation entre les locuteurs du Centre et du Sud (voir Annexe 26).

Pour les deux voyelles ouvertes /a ɑ/ et les voyelles médianes, la taille des ellipses est relativement grande, indiquant la grande variabilité inter-individuelle lors de leurs productions. Ainsi, nous présentons sur les Figures 101 à 103, les timbres vocaliques obtenus pour chaque locuteur. Malgré la variabilité dialectale et régionale observée en FR, toutes les voyelles se distinguent les unes des autres, sauf au niveau des deux voyelles ouvertes /a ɑ/ : il semblerait que ces deux voyelles ont fusionné pour ne former qu'une seule voyelle ouverte

⁶⁵ Il existe une différence significative entre les réalisations des locuteurs du Centre et du Sud de la France que nous ne traitons pas dans cette thèse. Pour plus de détails sur ces variations, voir Annexe 26, où nous avons présenté les réalisations individuelles des voyelles médianes en français en fonction des régions (Centre et Sud).

/a/, quels que soit le lieu d'articulation, la situation de production et le locuteur. Nous observons une variabilité inter-individuelle importante même en situation Isolation (voir Figure 103) : situation censé réduire les chevauchements entre les voyelles, car les voyelles produites dans cette situation de production sont proches des « cibles vocaliques ».

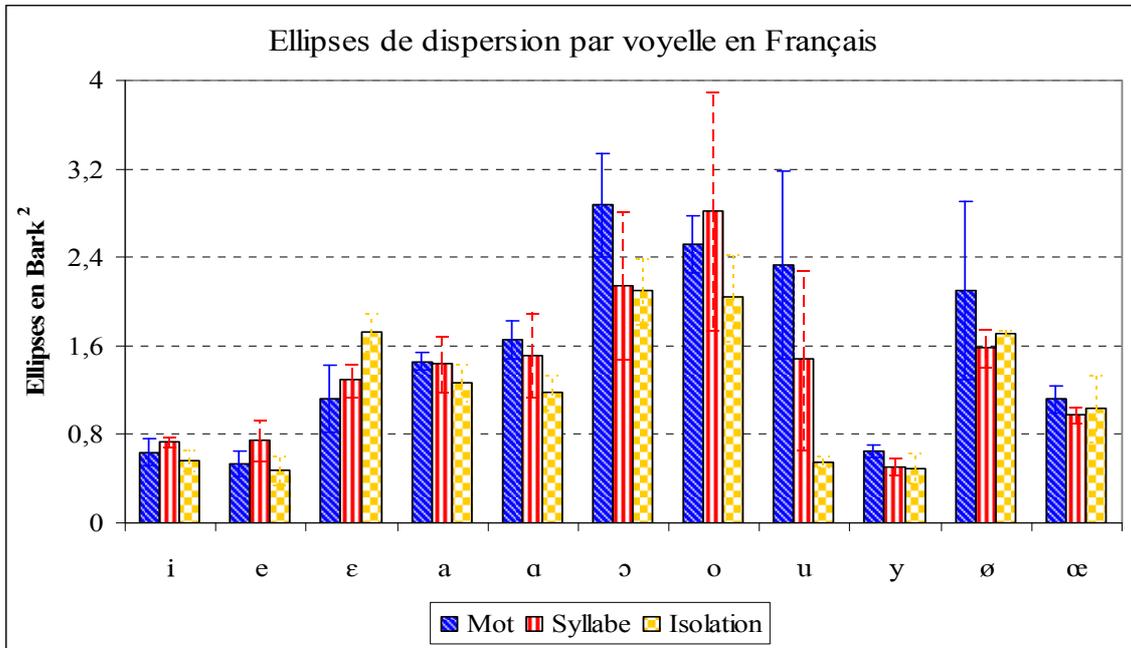


Figure 100 : Taille des ellipses vocaliques en FR en fonction de la situation de production.

Tableau 27 : Taille des ellipses vocaliques en fonction de la situation de production et du lieu d'articulation en FR (en Bark²).

(Bark²)		i	e	ε	a	ɑ	ɔ	o	u	y	ø	œ
Mot	b	0,51	0,64	1,32	1,53	1,55	3,08	2,32	3,31	0,58		1,25
	d	0,75	0,56	1,26	1,37	1,57	2,34	2,81	1,77	0,68	1,53	1,01
	k	0,67	0,41	0,77	1,47	1,85	3,21	2,42	1,92	0,68	2,67	1,09
Syllabe	b	0,79	0,95	1,46	1,24	1,42	2,33	2,02	1,25	0,53		1,01
	d	0,71	0,59	1,19	1,33	1,21	1,41	2,41	0,81	0,42	1,71	0,91
	k	0,72	0,71	1,22	1,73	1,93	2,71	4,05	2,38	0,58	1,47	1,04
Isolation	b	0,47	0,41	1,56	1,08	1,21	2,44	1,63	0,52	0,57		1,08
	d	0,56	0,41	1,72	1,38	1,02	1,97	2,42	0,52	0,34	1,68	1,31
	k	0,66	0,63	1,91	1,35	1,33	1,89	2,06	0,61	0,56	1,73	0,71

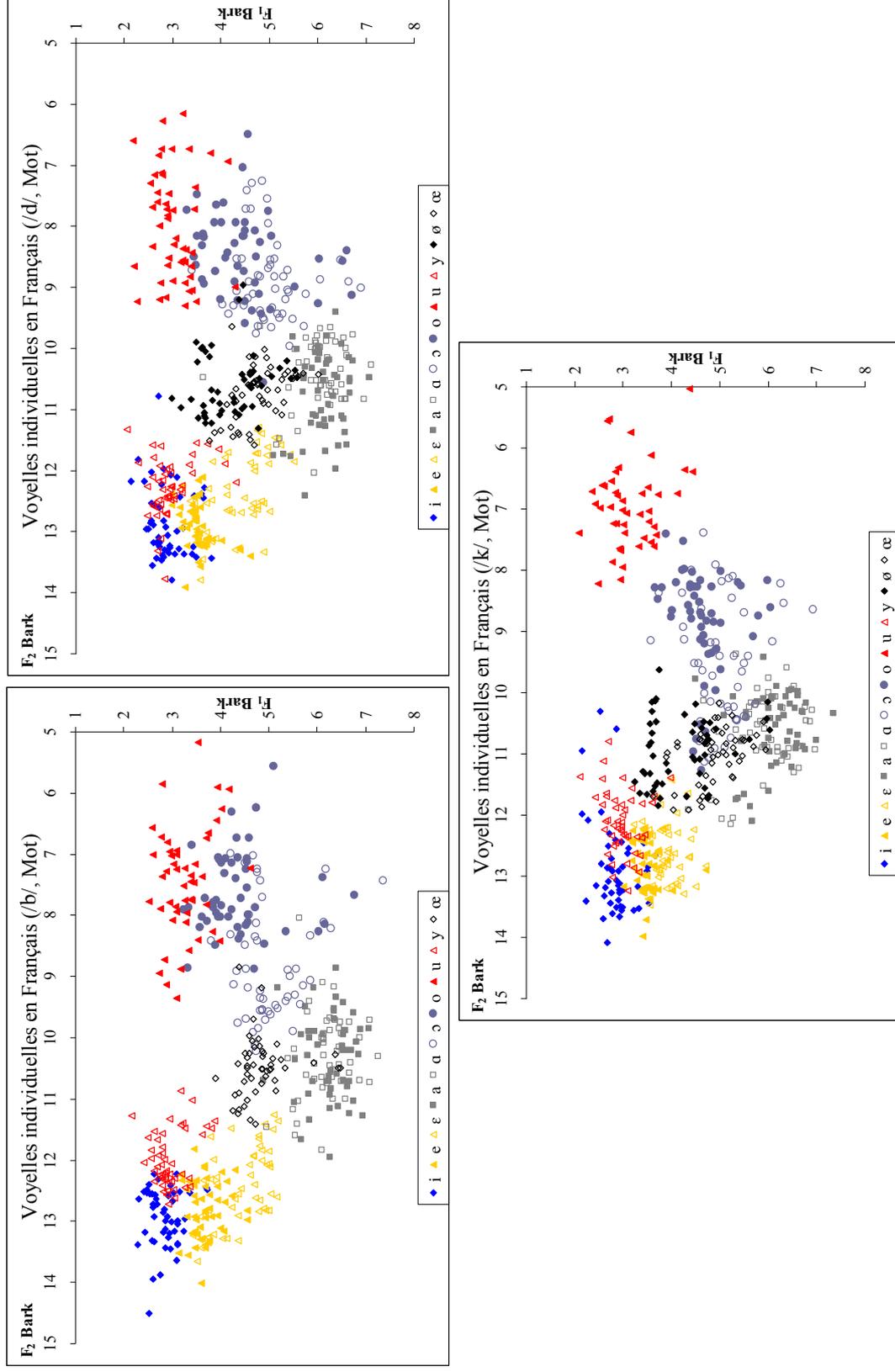


Figure 101 : Dispersion des voyelles en FR à partir des données individuelles, par lieu d'articulation, en Mot.

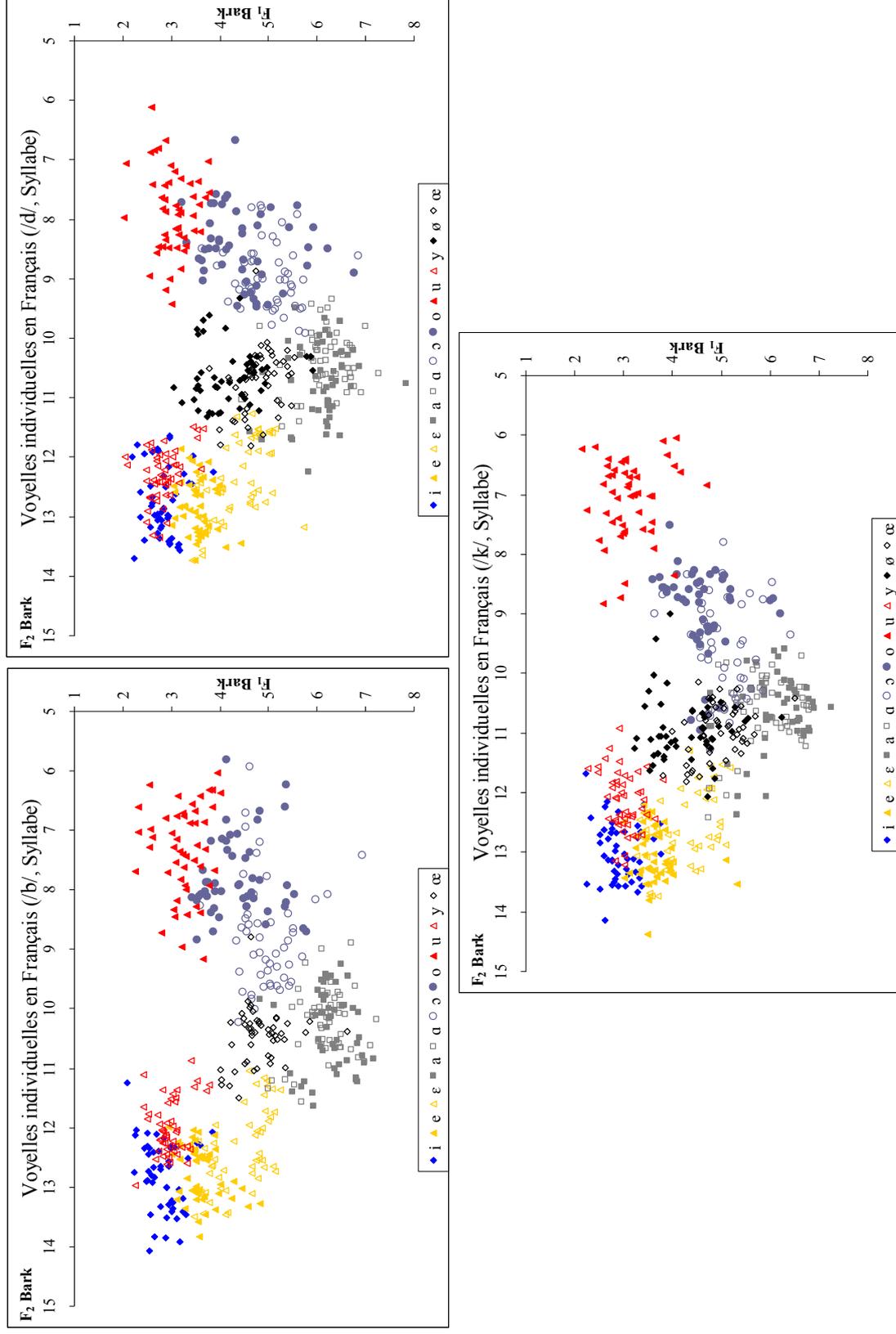


Figure 102 : Dispersion des voyelles en FR à partir des données individuelles, par lieu d'articulation, en Syllabe.

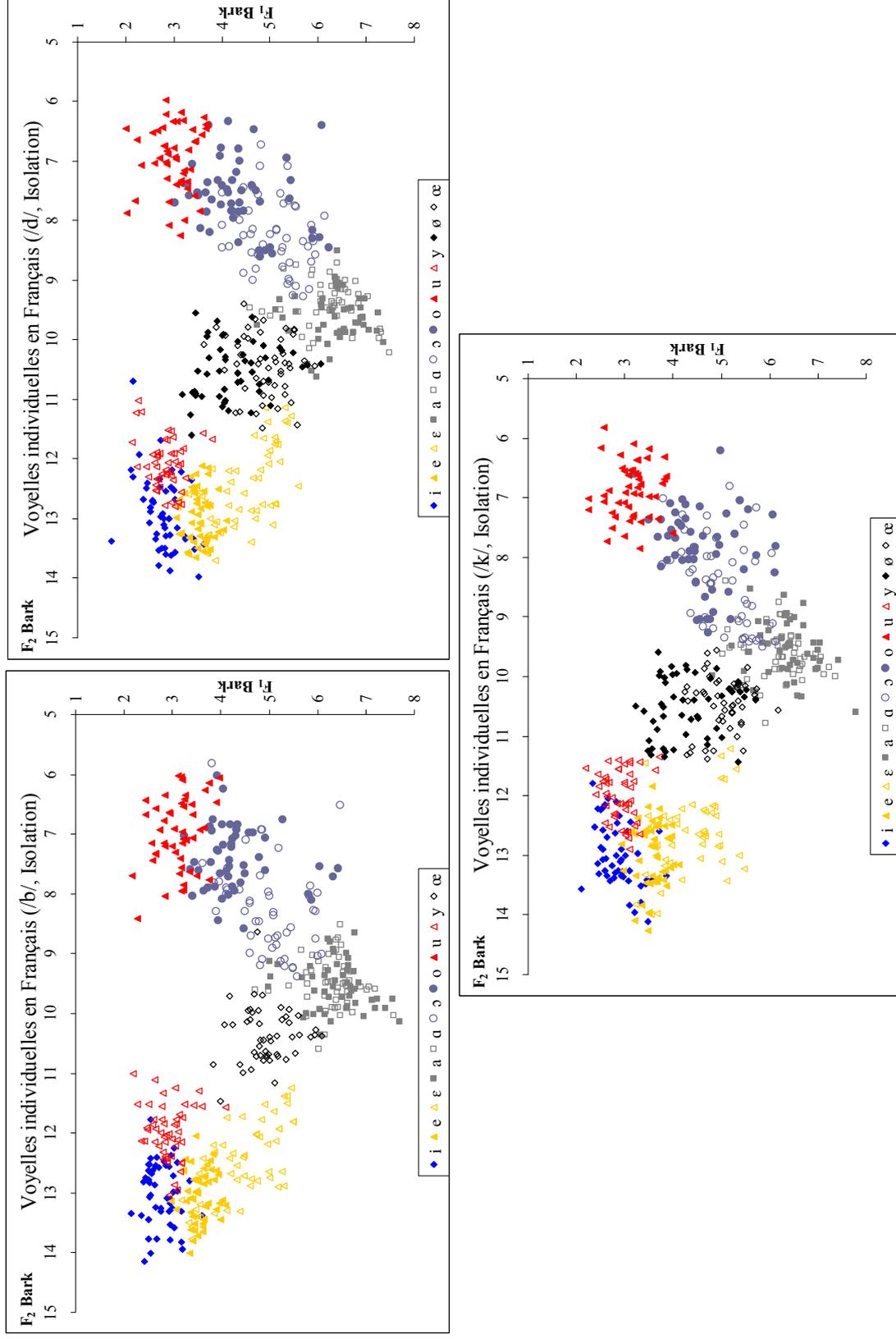


Figure 103 : Dispersion des voyelles en FR à partir des données individuelles, par lieu d'articulation, en Isolation.

4.4.3.2. Résultats en dynamique

Comme pour les deux autres langues, les résultats des **pent**es formantiques obtenues en FR, dans la situation de production Mot (voir Figures 104 à 106) montrent des différences importantes liées aux points de départ des transitions formantiques, à la cible vocalique dérivée et à leur degré d'inclinaison (pour les résultats en situation de production Syllabe et Isolation, voir Annexe 27).

Les résultats des pentes formantiques sur F_1 en FR (voir Figure 104) montrent des valeurs de l'**onset** vocalique différentes en fonction des timbres vocaliques et qui sont situées dans une fourchette de 2,8 à 4,5/5,2 Bark dans les environnements /b d k/ ($F(2, 1566) = 3,49$; $p < 0,05$). Ces différences sont concordant avec les degrés d'aperture des voyelles variables en FR : comme les valeurs de l'état stable de F_1 sont différentes en fonction des voyelles, celles-ci influencent directement celles de l'**onset** vocalique. Nous nous intéressons maintenant aux voyelles problématiques en FR, soit les voyelles « médianes » et les deux voyelles ouvertes /a α /. En ce qui concerne les voyelles « médianes » /e ϵ /-/ \emptyset $\text{\textcircled{a}}$ /-/ $\text{\textcircled{o}}$ $\text{\textcircled{u}}$ /, les valeurs de l'**onset** vocalique, de la cible vocalique dérivée et du degré d'inclinaison sont différentes en fonction du lieu d'articulation des consonnes adjacentes ($p < 0,001$). En contexte /b/, ces trois paramètres sont significativement différents entre les couples de voyelles /e ϵ / et / $\text{\textcircled{o}}$ $\text{\textcircled{u}}$ /, où le degré d'inclinaison des pentes formantiques est beaucoup plus marqué pour les dernières. En contexte /d/, les quatre voyelles /e \emptyset $\text{\textcircled{a}}$ $\text{\textcircled{o}}$ / présentent le même **onset** vocalique ; les trois voyelles / ϵ $\text{\textcircled{a}}$ $\text{\textcircled{o}}$ / obtiennent le même degré d'inclinaison ; et les couples de voyelles /e ϵ /-/ \emptyset $\text{\textcircled{a}}$ /-/ $\text{\textcircled{o}}$ $\text{\textcircled{u}}$ / sont significativement différents en fonction des trois paramètres. Enfin en contexte /k/, aucun des trois paramètres n'est homogène entre le couple de voyelles /e ϵ /-/ \emptyset $\text{\textcircled{a}}$ /-/ $\text{\textcircled{o}}$ $\text{\textcircled{u}}$ /, sauf pour le degré d'inclinaison qui est identique pour le couple /e ϵ /. En ce qui concerne les deux voyelles /a α /, nous obtenons des valeurs de l'**onset** vocalique, de la cible vocalique dérivée et du degré d'inclinaison différents dans les environnements /d k/ ; dans l'environnement /b/, les trois paramètres sont homogènes pour ces deux voyelles (effet non significatif).

Les valeurs des **pent**es formantiques, obtenues sur F_2 (voir Figure 105), sont significativement différentes entre les voyelles en FR et plus spécifiquement pour les voyelles problématiques. Ainsi, dans l'environnement /b/, les voyelles /e ε/-/o ɔ/ présentent des **onsets** et des cibles vocaliques dérivés et des degrés d'inclinaison des pentes différents. En contexte /d/, seuls les points de départ des transitions pour les couples /e ε/-/ø œ/-/o ɔ/ sont homogènes ; les valeurs de cible dérivée et de degré d'inclinaison sont différentes. Enfin dans l'environnement /d/, seul le couple /o ɔ/ présente des valeurs d'**onset**, de cibles vocaliques dérivées et de degrés d'inclinaison des pentes différentes. Les deux voyelles ouvertes /a α/ présentent des paramètres homogènes quel que soit le lieu d'articulation.

Quant aux résultats sur F_3 (voir Figure 106), les pentes formantiques obtenues sont significativement différentes dans les trois environnements consonantiques et pour les quatre couples de voyelles /e ε/-/ø œ/-/o ɔ/-/a α/, sauf pour le degré d'inclinaison des pentes qui est homogène pour les couples /o ɔ/-/a α/ en contexte /b/ ; /ø œ/-/o ɔ/-/a α/ en contexte /d/ et /o ɔ/ en contexte /k/.

Les résultats des **pent**es formantiques montrent que les voyelles « médianes » sont distinguables par leur **onset** et cible vocalique dérivée et par le degré d'inclinaison de leurs transitions formantiques. Pour ce qui est des deux voyelles ouvertes /a α/, celles-ci semblent fusionner pour n'en former qu'une seule ; résultats allant dans le sens de ceux obtenus pour l'analyse **statique** des voyelles. Nous nous intéressons dans la partie qui suit aux résultats des classifications des voyelles en FR à partir d'une analyse discriminante.

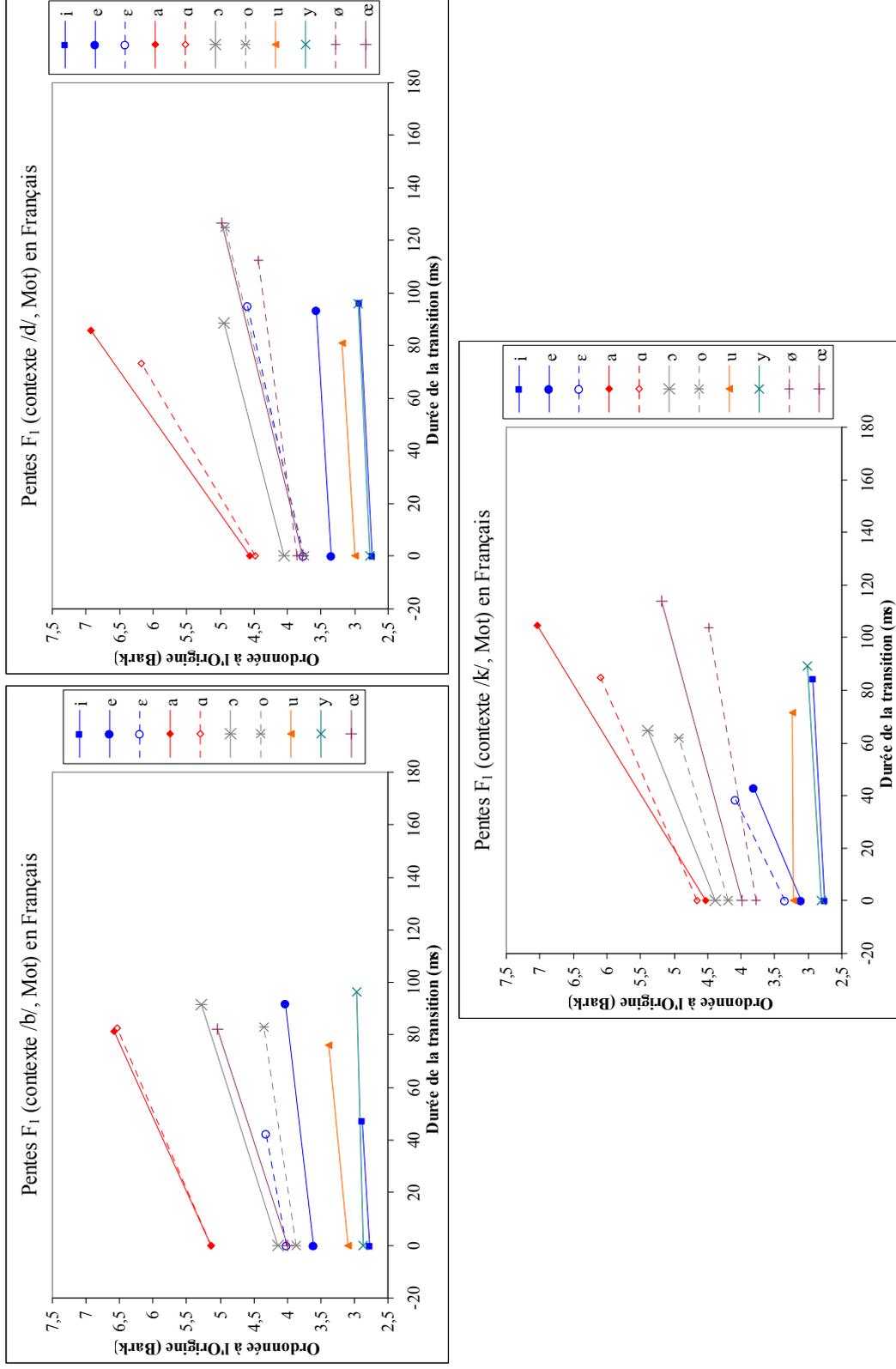


Figure 104 : Pentas formantiques de F₁ des voyelles en FR sans normalisation du temps, en fonction du lieu d'articulation en Mot.

RÉSULTATS EN PRODUCTION DES VOYELLES

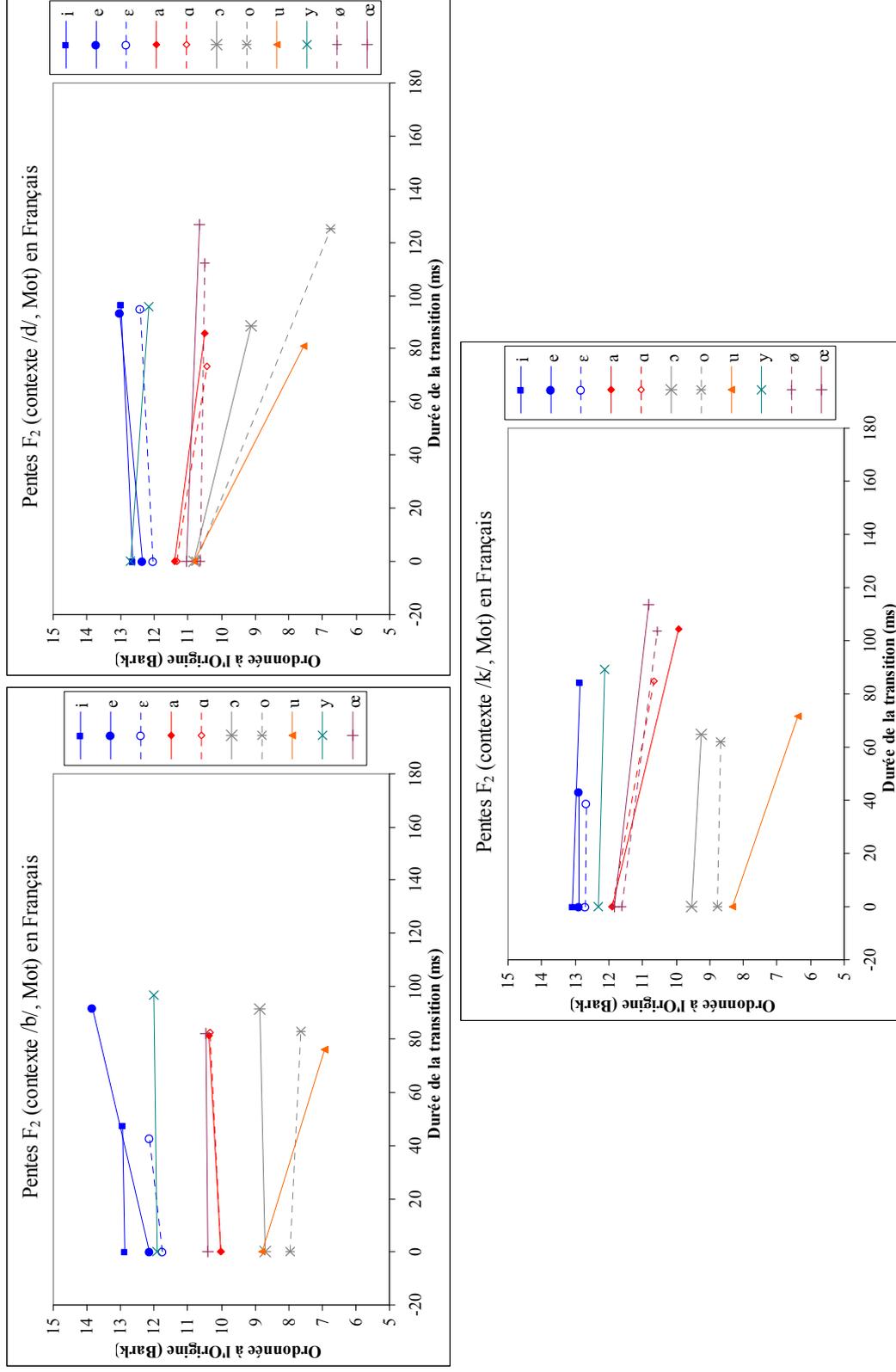


Figure 105 : Pentas formantiques de F₂ des voyelles en FR sans normalisation du temps, en fonction du lieu d'articulation en Mot.

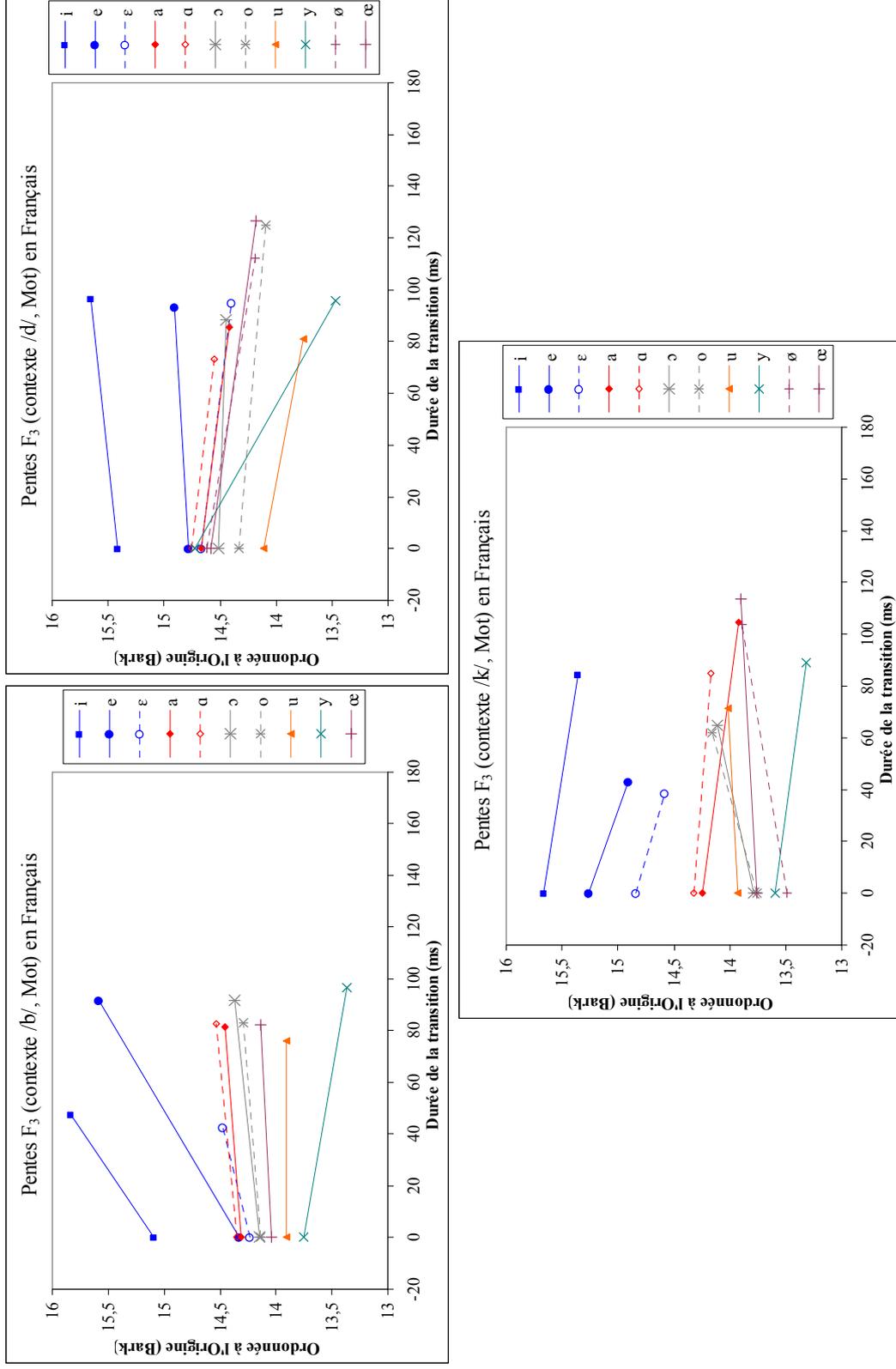


Figure 106 : Pentes formantiques de F₃ des voyelles en FR sans normalisation du temps, en fonction du lieu d'articulation en Mot.

Les résultats obtenus dans ces parties ont montré que les voyelles « médianes » du FR /e ε ø œ o ɔ/ sont statistiquement différentes du point de vue des analyses **statiques** et **dynamiques** : les « cibles vocaliques » sont significativement différentes quels que soient la situation de production, le lieu d'articulation et le locuteur ; les **pentés formantiques** de ces voyelles quant à elles, sont significativement différentes en ce qui concerne le point de départ des transitions formantiques, le degré d'inclinaison des transitions et la position des « cibles vocaliques » dérivées. En ce qui concerne les deux voyelles ouvertes /a α/, nos résultats ont montré qu'elles constituent une seule et unique voyelle /a/ en **statique** et en **dynamique** et ce quels que soient la situation de production, le lieu d'articulation et le locuteur. Nous nous attendons donc à observer le maximum de confusion lors des analyses discriminante entre les deux voyelles ouvertes et peu voire pas du tout de confusion entre les coupes des voyelles « médianes » /e ε/, /ø œ/ et /o ɔ/, résultats présentés dans la partie qui suit.

4.4.3.3. Résultats de l'analyse discriminante

4.4.3.3.1. En statique

Une analyse discriminante, avec validation croisée ayant comme paramètres les valeurs formantiques de F_1 , F_2 & F_3 des voyelles du FR, en incluant et en excluant la durée vocalique, a été appliquée. Les taux de classification correcte des voyelles en fonction des situations de production et des lieux d'articulation sont présentés dans le Tableau 28.

Tableau 28 : Taux de classification correcte des voyelles en FR avec les valeurs formantiques de F_1 , F_2 & F_3 en fonction de la situation de production et du lieu d'articulation. Les valeurs entre parenthèses indiquent les taux de classification en excluant la durée. (Taux significatifs Box's M ; $p < 0,001$).

	/b/	/d/	/k/
Mot	73,4% (70,6%)	72,5% (68,0%)	72,7% (64,7%)
Syllabe	69,9% (71,5%)	76,4% (72,0%)	70,0% (64,0%)
Isolation	74,2% (74,6%)	71,6% (71,5%)	66,7% (68,0%)

Les taux obtenus sont moyennement élevés. Lorsque le paramètre de la durée vocalique est inclus dans l'analyse discriminante, le taux de classification correcte est augmenté en moyenne de 2,5% (71,9% avec durée contre 69,4% sans durée), (BinomialQ ;

$p < 0,01$). On sait qu'en FR la distinction de durée vocalique n'est pas pertinente, mais les résultats de l'analyse discriminante montrent qu'en situation Mot, le taux de classification a augmenté de 5% en moyenne, tandis qu'en Isolation, l'utilisation de la durée n'a pas amélioré la classification. Systématiquement, lorsque le paramètre de la durée vocalique est ajouté à l'analyse, celui-ci était toujours le premier à être utilisé comme variable classificatoire. Toutes les confusions observées dans la classification des voyelles du FR, sont dues aux voyelles « problématiques », où un chevauchement important est observé.

Les résultats graphiques des confusions dans la classification des voyelles du FR en situation Mot sont présentés sur les Figures 107 et 108 (pour les résultats graphiques en situation Syllabe et Isolation, voir les Annexes 28 et 29). Nous obtenons approximativement les mêmes taux de confusion entre les différentes voyelles lorsque le paramètre de la durée vocalique est inclus ou exclu.

En ce qui concerne les deux voyelles /a a/, les résultats graphiques montrent qu'elles sont tout le temps confondues entre elles, sauf dans l'environnement /k/ avec la durée vocalique ajouté (voir Figure 107), où /a/ est classifiée correctement à plus de 80% et /a/, à plus de 60%. Les voyelles /e ε/ sont toujours les seules à être bien classifiées avec des taux dépassant les 80% (en plus des voyelles /i u y/); les autres voyelles « médianes » sont souvent confondues entre elles.

Ces confusions importantes entre d'un côté les deux voyelles ouvertes /a a/ et de l'autre entre les voyelles « médianes » /e ε ø œ o ɔ/, vont dans le sens de nos résultats. Vu la proximité des deux voyelles ouvertes, elles étaient toujours confondues entre elles, les autres voyelles « médianes » montrent des profils différents.

Nous nous intéressons dans les parties qui suivent aux résultats des différentes analyses discriminantes appliquées sur les données en **dynamique**.

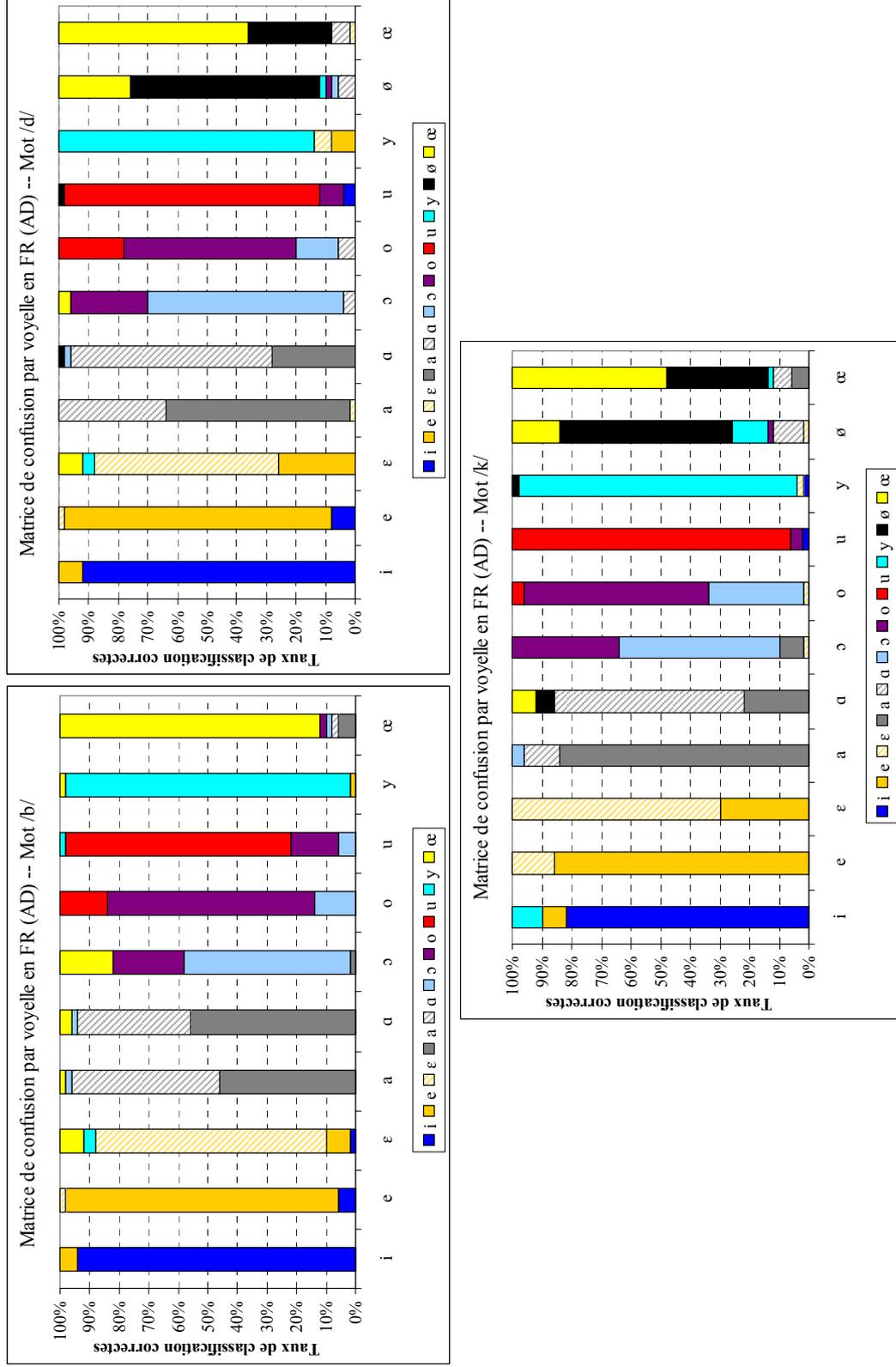


Figure 107 : Matrices de confusion de classification des voyelles du FR en **statique** avec la durée vocalique (AD) en situation de production Mot et en fonction du lieu d'articulation.

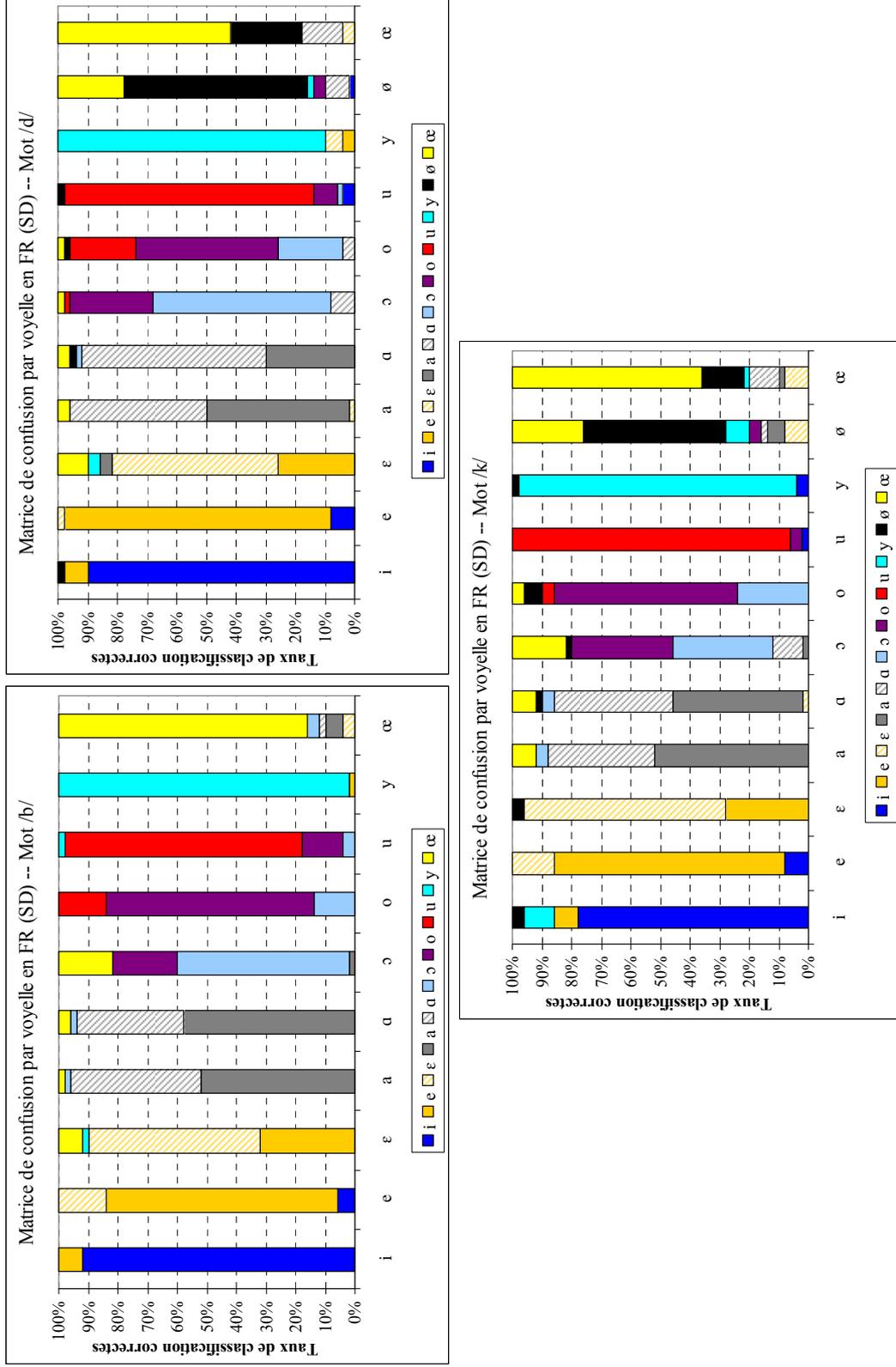


Figure 108 : Matrices de confusion de classification des voyelles du FR en statique sans la durée vocalique (SD) en situation de production Mot et en fonction du lieu d'articulation.

4.4.3.3.2. En dynamique

4.4.3.3.2.1. Régression linéaire

Les paramètres utilisés pour cette analyse sont les valeurs des **pentés formantiques** et des ordonnées à l'origine de F_1 , F_2 & F_3 des voyelles du FR, ainsi que la durée de la transition formantique. Les **pentés formantiques** obtenues en fonction des voyelles, des situations de production et des lieux d'articulation sont significativement différentes. L'analyse discriminante a permis de séparer les voyelles en FR, dont les taux sont présentés dans le Tableau 29.

Tableau 29 : Taux de classification correcte des voyelles du FR à partir de la régression linéaire en fonction de la situation de production et du lieu d'articulation. (Taux significatifs, Box's M ; $p < 0,001$).

	/b/	/d/	/k/
Mot	76,6%	74,2%	71,0%
Syllabe	73,5%	77,1%	72,7%
Isolation	75,0%	72,2%	67,1%

Les coefficients de la régression linéaire permettent de discriminer les voyelles en FR en fonction de la situation de production et du lieu d'articulation des consonnes. Toutefois, l'examen détaillé des taux de classification de l'analyse discriminante révèle toujours des confusions de classification liées aux voyelles problématiques en FR : les voyelles « médianes » /e ε/-/ø œ/-/o ɔ/ et les voyelles ouvertes /a α/. Nous présentons sur la Figure 109 les taux de confusion dans la classification des voyelles du FR en situation de production Mot (pour les deux autres situations de production, voir Annexe 30).

Les taux de confusion montrent globalement que les trois voyelles /i y u/ ont été classifiées correctement avec des taux supérieurs à 85%. Les faibles confusions obtenues pour les voyelles « médianes » montrent globalement que ces voyelles restent distinctes les unes des autres, i.e., pas de fusion entre les catégories vocaliques. En revanche, les taux de confusion des voyelles ouvertes montrent que la production de celles-ci est variable en fonction des locuteurs français et que la fusion de ces deux voyelles est bien présente.

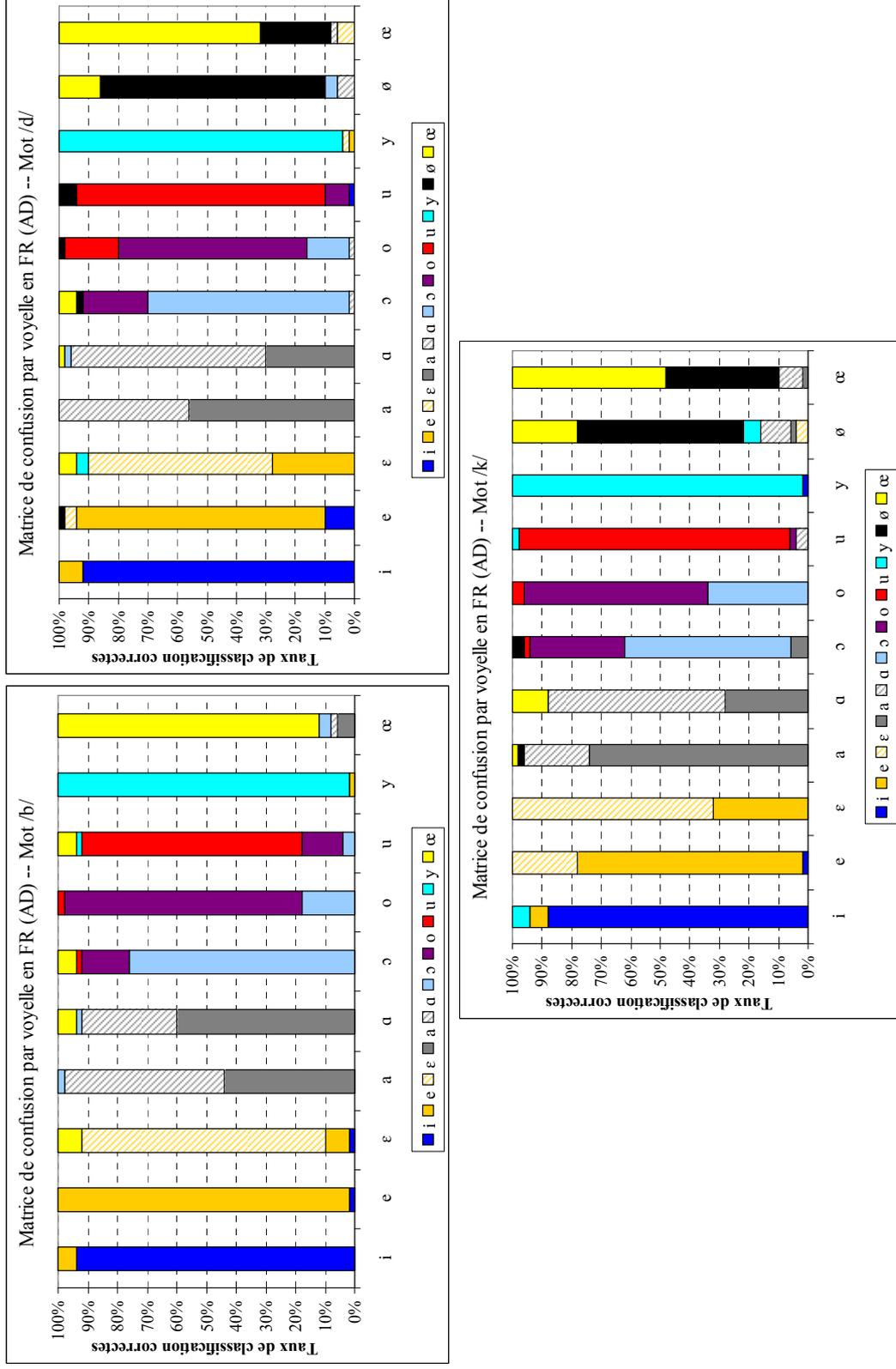


Figure 109 : Matrices de confusion de classification des voyelles du FR en dynamique (régression linéaire) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

La prochaine étape consiste à évaluer l'impact de l'analyse **dynamique** non linéaire des voyelles par une régression polynomiales (de 2^{ème} et 3^{ème} ordre) sur les taux de classification des voyelles du FR.

4.4.3.3.2.2. Régressions polynomiales de 2^{ème} et 3^{ème} ordre

Les paramètres utilisés pour cette analyse sont tous les coefficients des régressions polynomiales de 2^{ème} et de 3^{ème} ordre⁶⁶, ainsi que la durée de la transition formantique. Les résultats de la classification des voyelles du FR en fonction de la situation de production et du lieu d'articulation sont présentés dans le Tableau 30.

Tableau 30 : Taux de classification correcte des voyelles du FR à partir des régressions polynomiales de 2^{ème} et de 3^{ème} ordre, en fonction de la situation de production et du lieu d'articulation. Les valeurs entre parenthèses indiquent les taux de classification avec la régression polynomiale de 3^{ème} ordre. (Taux sont Box's M ; $p < 0,001$).

	/b/	/d/	/k/
Mot	74,4% (73,2%)	73,8% (74,2%)	72,1% (71,6%)
Syllabe	73,1% (75,4%)	75,1% (74,7%)	70,4% (70,5%)
Isolation	74,6% (74,2%)	71,3% (70,2%)	68,5% (67,8%)

Ces résultats montrent la possibilité de discriminer les voyelles en FR en fonction de la situation de production et le lieu d'articulation avec les coefficients de régressions polynomiales de 2nd et 3^{ème} ordre, des trois premiers formants ainsi que la durée de la transition formantique. Nous n'obtenons pas d'amélioration du taux de classification correcte en utilisant le 3^{ème} ordre de la régression polynomiale en comparaison avec le 2^{ème} ordre. En moyenne, le taux de classification correcte baisse de 0,2% en utilisant la régression polynomiale de 3^{ème} ordre (72,6% avec la régression polynomiale de 2^{ème} ordre contre 72,4% avec le 3^{ème} ordre), (BinomialQ ; $p=1$). Comme pour l'AM et l'AJ, l'augmentation de l'ordre de la régression polynomiale n'améliore pas la caractérisation des voyelles du FR.

⁶⁶ Les coefficients utilisés sont : pour la régression polynomiale de 2^{ème} ordre, a_0 = la moyenne des valeurs, a_1 = la droite de régression et a_2 = l'ordonnée à l'origine ; pour la régression polynomiale de 3^{ème} ordre, a_0 = la moyenne des valeurs, a_1 = la droite de régression, a_2 = l'excursion parabolique de la voyelle indépendamment de sa droite et a_3 = l'ordonnée à l'origine.

Nous obtenons des taux de classification en moyenne bas avec les deux analyses de régression polynomiale de 2^{ème} et de 3^{ème} ordre. Les confusions importantes dans la classification sont à l'origine de cette baisse des taux. Nous présentons sur les Figures 110 et 111 ces confusions dans la situation de production Mot (pour les deux autres situations, voir Annexes 31 et 32). Les voyelles les mieux classifiées sont toujours les trois voyelles fermées /i y u/, car elles se trouvent dans l'extrémité haute de l'espace vocalique et aucune autre voyelle n'est à proximité directe avec l'une d'entre elles. Nous obtenons une homogénéité dans les confusions de classification quels que soient la situation de production et l'analyse de régression entreprise.

Les erreurs de classification en ce qui concerne les voyelles « problématiques » montrent que les deux voyelles /a ɑ/ sont toujours confondues entre elles, indiquant que la structure de leurs trajectoires formantiques est identique. Pour ce qui est des couples de voyelles /e ε/, /ø œ/ et /o ɔ/, les résultats montrent que c'est toujours le dernier couple de voyelles qui le plus confondu, en revanche les deux voyelles /e ε/, sont toujours bien classifiées avec des taux proches des 80%.

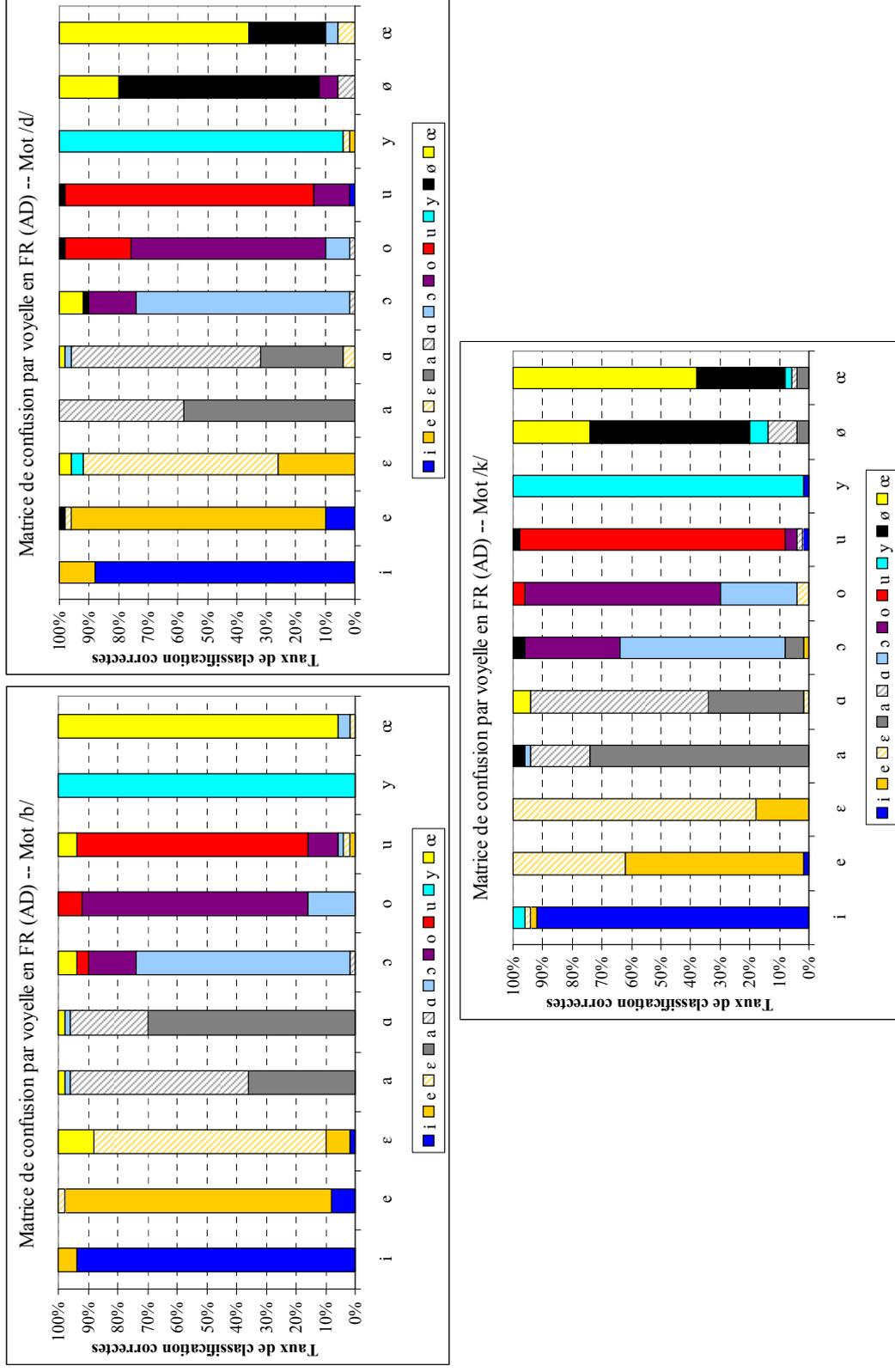


Figure 110 : Matrices de confusion de classification des voyelles du FR en dynamique (régression polynomiale de 2^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

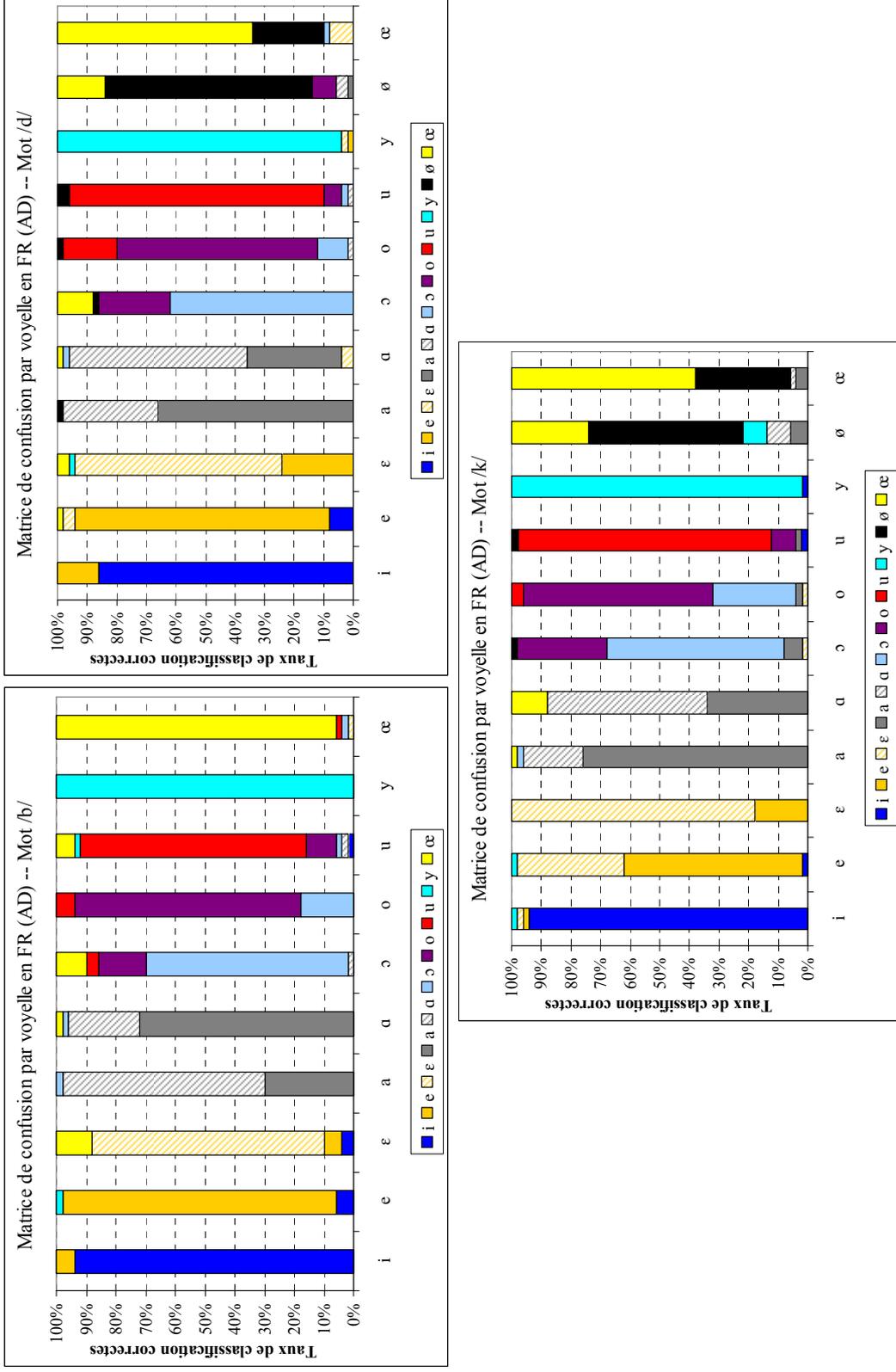


Figure 111 : Matrices de confusion de classification des voyelles du FR en dynamique (régression polynomiale de 3^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

4.4.3.4. Discussion des résultats en statique vs en dynamique

L'examen détaillé des taux d'erreur de classification de l'analyse discriminante effectuée en **statique** (à partir des « cibles vocaliques ») et en **dynamique** (à partir des deux analyses de régressions linéaire et polynomiale), montre que les voyelles « problématiques » du FR ont été toujours confondues entre elles. Nous avons montré dans les parties précédentes qu'avec les analyses **statiques** des voyelles en FR (voir § 4.4.3), certaines voyelles médianes sont très rapprochées et ce à cause de l'absence de distinction par les locuteurs du Sud des voyelles mi-fermées vs mi-ouvertes. Pour les deux voyelles ouvertes /a a/, nous observons l'absence de distinction par les locuteurs FR de celles-ci. Les confusions dans la classification de ces deux voyelles confortent l'idée selon laquelle elles ne constituent qu'une seule et unique voyelle. Du point de vue des analyses **dynamiques**, nos résultats ont montré que les trajectoires formantiques des voyelles « problématiques » sont significativement différentes. En effet, nous obtenons des différences entre les degrés d'inclinaison des transitions formantiques, les points de départ des transitions et les « cibles vocaliques » dérivées, pour les voyelles « médianes », en ce qui concerne les deux voyelles ouvertes, elles présentent les mêmes structures **dynamiques**.

Nous avons effectué une comparaison entre les taux de classification correcte obtenus à partir des différentes analyses discriminante et dont les résultats sont présentés sur la Figure 112. La comparaison des taux obtenus en situation Mot montre que l'utilisation de n'importe quelle analyse **dynamique** (i.e., régressions linéaire et polynomiales de 2^{ème} et de 3^{ème} ordre), mais également l'utilisation des « cibles vocaliques » avec la durée vocalique, augmente le taux moyen significativement de 2 à 4% par rapport à l'utilisation des paramètres **statiques** (BinomialQ, $p < 0,001$). Les taux moyens obtenus dans les trois analyses **dynamiques** en comparaison avec ceux obtenus à partir des « cibles vocaliques » avec la durée vocalique sont significativement non différentes. Nous n'observons aucune amélioration du taux de classification pour la comparaison entre ces différents paramètres « **dynamiques** ». Nous nous attendions à obtenir une augmentation des taux de classification avec les analyses de

régression, car les trajectoires formantiques sont mieux modélisées. Il semblerait donc que la durée vocalique est un facteur discriminant, durée non pertinente dans la description des voyelles du FR. En situation Syllabe, nous obtenons toujours une augmentation moyenne significative des taux de classification avec tous les paramètres « **dynamiques** » (les trois analyses de régression et les « cibles vocaliques » avec la durée vocalique) de 3 à 5% par rapport au taux obtenu pour les analyses **statiques** (BinomialQ ; $p < 0,001$). En situation Isolation, nous n'obtenons aucune différence significative entre les taux obtenus pour les analyses **statiques** et **dynamiques**. En effet, ce sont les différences de « cibles vocaliques » obtenues pour les voyelles du FR produites dans cette situation de production qui expliquent l'homogénéité des résultats. Les locuteurs français produisent des voyelles soutenues en Isolation, i.e., sans modification des structures formantiques **intrinsèques** des voyelles.

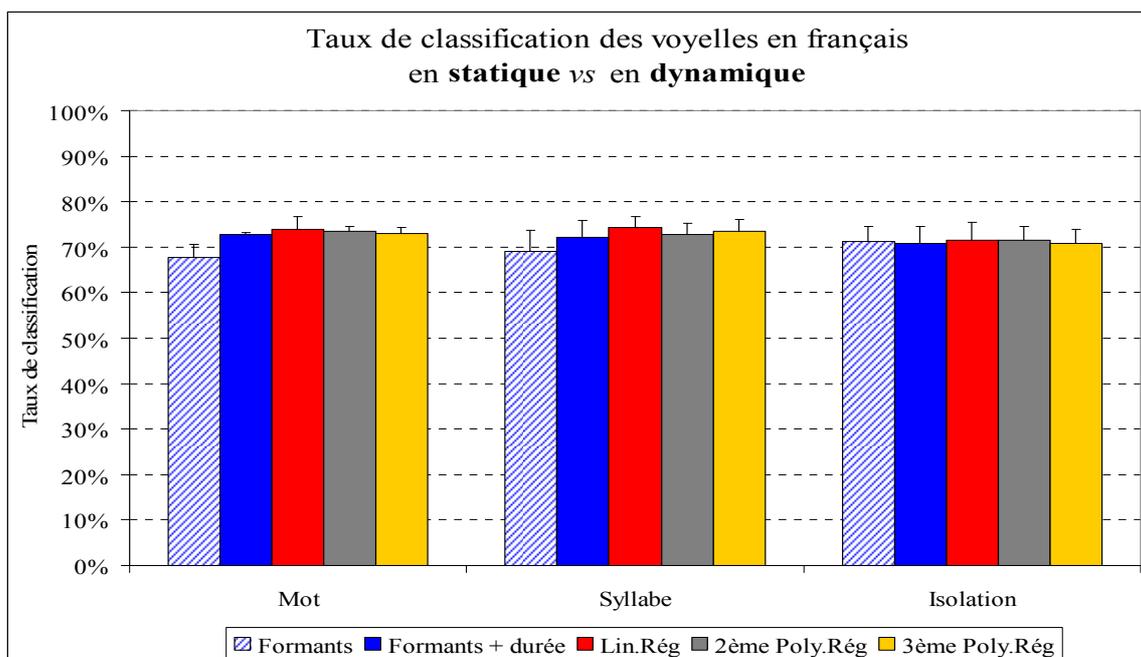


Figure 112 : Taux de classification correcte moyens des voyelles du FR produites dans les contextes /b d k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur).
 Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} ou 3^{ème} ordre).

Les différents résultats que nous avons présentés dans les parties précédentes ont montré, d'un côté, que les deux voyelles ouvertes /a a/ sont réalisées identiquement, de l'autre, que les voyelles « médianes » étaient produites différemment par les locuteurs français, avec néanmoins des chevauchements entre les différentes catégories vocaliques. Ainsi, nous proposons le système vocalique du FR comme ayant les dix voyelles orales

suivantes /i e ε a ɔ o u y ø œ/, sans la prise en compte des variations régionales affectant les voyelles dites « médianes ».

4.5. RÉCAPITULATIF DES RÉSULTATS EN STATIQUE VS EN DYNAMIQUE

La comparaison entre les trois langues, i.e., avec des systèmes de densité différente, montre des différences significatives dans l'organisation de leurs espaces vocaliques. En effet, les résultats en **statique** ont révélés une aire de dispersion obtenue pour les voyelles /i a u/ (voyelles longues en arabe) plus petite en AM par rapport à celle en FR qui obtient l'aire de dispersion la plus grande ; la différence entre les aires obtenues en AM et en AJ n'est pas significatives. Lorsque la comparaison entre les trois langues s'effectue à partir de toutes les voyelles, les résultats montrent que l'aire de dispersion en FR est toujours la plus grande significativement par rapport à celle en AM ou en AJ, mais également que l'aire obtenue en AJ est significativement plus grande que celle de l'AM. En ce qui concerne l'effet de la densité des systèmes sur la taille des ellipses de dispersion de chaque voyelle, nos résultats montrent l'effet inverse, i.e., des aires de dispersion des catégories vocaliques /i a u/ plus grandes en AJ par rapport à celles en FR ; l'AM quant à lui obtient les aires les plus petites. Dans les parties précédentes, nous avons mis en évidence les différentes théories, la **théorie quantique (TQ)**, la **théorie de la dispersion adaptée (TDA)** et la **théorie de la dispersion-focalisation (TDF)**, qui ont proposé quelques principes universels qui visent à expliquer les effets de la densité des systèmes sur l'organisation des voyelles dans les langues du monde. Nous rappelons les quatre cas présentés sur la Figure 26, p. 113 :

- Cas 1 L'espace vocalique n'est pas fonction de la densité des systèmes et la taille des ellipses vocaliques est identique entre les langues, suivant les propositions de la **TQ** et de la **TDF**,
- Cas 2 L'espace vocalique est fonction de la densité des systèmes et la taille des ellipses vocaliques est identique entre les langues suivant les propositions de la **TDA**,
- Cas 3 L'espace vocalique entre les langues n'est pas fonction de la densité des systèmes et la taille des ellipses vocaliques dépend du nombre de segment vocalique, suivant les propositions de la **TQ**, de la **TDA** et de la **TDF**,

Cas 4 L'espace vocalique et la taille des ellipses vocalique entre les langues sont fonction de la densité des systèmes, suivant les propositions de la **TDA**,

Ces constats vont dans le sens de l'hypothèse d'une influence de la densité des systèmes vocaliques sur l'organisation des espaces vocaliques dans les trois langues étudiées. l'AM semble respecter les propositions décrites dans le Cas N°. 2 ; l'AJ et le FR, le Cas N°. 3 : un espace vocalique « très » réduit avec des ellipses de dispersion vocalique réduites (le cas de l'AM) ; et un espace vocalique plus grand avec une dispersion intra-catégorie vocalique grande (le cas de l'AJ et du FR).

Les résultats obtenus en **dynamique**, montrent des **pentés formantiques** moins importantes sur F_1 , F_2 & F_3 en AM par rapport à celles en AJ ; le FR obtient les trajectoires de formants les plus importants, avec et sans normalisation du temps. Ces résultats sont en accord avec les prédictions que nous avons formulées. Nous observons en effet que pour une même voyelle, produite dans un même environnement consonantique et dans la même situation de production, les **pentés formantiques** obtenues pour un formant précis sont plus « raides » pour les langues présentant un système vocalique à forte densité en comparaison avec celles ayant un système vocalique à faible densité. L'explication que nous avons avancée est la suivante : plus une langue a des voyelles, plus ces voyelles doivent être significativement distinctes et ainsi leurs productions se rapprocheraient de celles des voyelles **prototypiques**. Autrement dit, l'effet de « périphérisation » de l'espace vocalique causé par le nombre plus grand de voyelles affecte la position de toutes les voyelles et plus particulièrement les voyelles « extrêmes », qui à leurs tours affectent les points de départ des transitions formantiques : effet qui se traduit directement sur le degré d'inclinaison des **pentés formantiques**.

À partir de la comparaison inter-langues effectuée à partir des analyses **statiques** et **dynamiques**, nos résultats ont montré que la densité des systèmes vocaliques semble affecter l'organisation de toutes les voyelles dans chaque langue. Les « cibles vocaliques » obtenues dans les trois langues sont *Hyper-Articulées* en FR et *Hypo-Articulées* en AM ; l'AJ se situe entre les deux. Les **pentés formantiques** obtenues en AM sont plus « aplaties » que celles obtenues en AJ, en FR en revanche, elles sont plus « raides ». L'*Hyper-Articulation* des

voyelles du FR affecte directement les points de départ des voyelles et par conséquent le degré d'inclinaison des transitions formantiques.

En ce qui concerne l'évaluation de la spécificité de chaque langue, les analyses **statiques** des voyelles ont montré des dispersions des voyelles en moyenne plus grandes en situation de production Isolation par rapport à celles des mêmes voyelles produites en Syllabe ou en Mot, en AM et FR ; aucune différence significative n'a été obtenue en AJ. Ceci peut s'expliquer par la difficulté qu'avaient les locuteurs jordaniens à produire les voyelles en Syllabe ou Isolation : ils tendaient à produire la plupart des voyelles comme une voyelle [a] ou [ə] ; certains locuteurs produisaient ces mêmes voyelles beaucoup plus « centralisées » que leurs correspondantes en Mot et/ou en Syllabe.

Pour les trois langues, le lieu d'articulation des consonnes adjacentes affecte les timbres des voyelles, même lorsqu'elles sont produites en contexte Isolation : *la tâche des locuteurs consistait à produire les voyelles comme réalisées en Isolation en respectant la prononciation exacte de celles-ci dans situation de production Mot*. La dispersion de chaque catégorie vocalique était plus petite en Mot et en Syllabe par rapport à celle en contexte Isolation en AM seulement ; aucun effet n'a été obtenu pour l'AJ ou le FR. Nous nous attendions à observer l'effet inverse en AM ; i.e., une dispersion de chaque catégorie vocalique moins importante en Isolation, car la précision articulatoire lors de la production des voyelles en Isolation affecte la position des voyelles produites : les voyelles « périphériques » se trouveraient dans les extrémités de l'espace vocalique afin de garantir une distinction suffisante entre elles et entraîne par conséquent une production quasi-identique des timbres vocaliques quels que soient la répétition et le locuteur.

Quant aux analyses **dynamiques**, les résultats au niveau intra-langue montrent un effet significatif des trajectoires de formants qui sont plus importants en Mot et en Syllabe comparées à celles produites en Isolation dans les trois langues. Nous n'avons pas observé d'effet du lieu d'articulation vélaire sur les mouvements des formants : pour F_1 , les trajectoires de formants sont plus importantes dans l'environnement /k/ par rapport aux bilabial ou dental en FR uniquement, pour F_2 , et dans les trois langues, un effet significatif

des mouvements de formants en contexte vélaire a été observé, et enfin pour F_3 , aucun effet du lieu d'articulation vélaire n'a été obtenu.

À partir des résultats obtenus par les **équations de locus** pour F_2 (voir § 4.3.2), la littérature propose une quantification du degré de coarticulation entre consonne et voyelle : le degré de coarticulation est plus important en contexte vélaire, intermédiaire en contexte bilabial et moins important en contexte dental. Les résultats que nous avons obtenus en **dynamique**, à partir des **pentés formantiques**, sont conformes. À l'inverse des mesures proposées par les **équations de locus**, nous pouvons quantifier le degré de coarticulation entre consonnes et voyelles pour chaque voyelle séparément et pour chaque formant vocalique avec la méthode que nous avons proposée.

Nos résultats montrent également que quels que soit la langue et le lieu d'articulation, les trajectoires des formants obtenues sur F_1 , sont les plus importantes pour les deux voyelles ouvertes /a a/ par rapport à /i/ ou /u/, les voyelles médianes obtiennent des trajectoires de formants intermédiaires ; sur F_2 , en revanche, /u/ et les voyelles postérieures obtiennent les trajectoires de formants les plus importantes, /i/ et les voyelles antérieures les moins importants et /a/ et les voyelles centrales entre les deux (sauf en AM). Les trajectoires de formants obtenues correspondent à des modifications des valeurs formantiques des voyelles à l'état stable. Lorsqu'on obtient des trajectoires de formants importantes, une variation importante entre la valeur formantique à l'**onset** et au milieu temporel de la voyelle est observée (i.e., une centralisation sur F_1 et/ou F_2).

Les résultats que nous avons obtenus dans la comparaison inter-langues et intra-langue montrent approximativement les mêmes tendances : la **dynamique** apporte plus de précision à la quantification des trajectoires formantiques obtenues dans les trois situations de production, résultats obtenus à partir des taux moyens de classification des analyses discriminantes (sauf en FR où l'utilisation de la **dynamique** n'a pas amélioré les taux de classification moyens dans la situation Isolation).

Les taux de classification correcte ont augmenté en moyenne de 10 à 30 % (selon le cas) en utilisant les différentes analyses **dynamiques** pour les comparaisons intra- et inter-langues. L'amélioration du taux la plus marquante est dans la classification des trois langues

en fonction de la situation de production et du lieu d'articulation en **statique** vs **dynamique** (39,3% avec les valeurs centrales des formants contre 64,1% avec la régression polynomiale de 3^{ème} ordre), (BinomialQ ; $p < 0,001$). Nous obtenons une amélioration du taux de classification et ce dès l'ajout de la durée de la transition formantique en moyenne de 15 à 20%, (BinomialQ ; $p < 0,001$). Nous observons une baisse des taux de classification moyens en normalisant le temps, mais ils restent néanmoins supérieurs aux taux de classification en utilisant les valeurs formantiques à l'état stable uniquement.

L'utilisation des différentes analyses de régression (linéaire et polynomiales de 2^{ème} et de 3^{ème} ordre) a permis de mettre en évidence les trajectoires formantiques **intrinsèques** et **extrinsèques** dans chaque langues. Les trajectoires formantiques **intrinsèques** ont été surtout quantifiées dans la situation de production Isolation, car ce sont les changements formantiques propres aux voyelles qui ont été modélisés. Les trajectoires formantiques **extrinsèques** ont été surtout quantifiées dans les situations de production Mot et Syllabe, où les effets coarticulatoires des consonnes adjacentes ont été modélisés. L'utilisation de l'analyse de régression linéaire a permis quelques fois de bien modéliser les trajectoires formantiques, toutefois, il a été nécessaire d'utiliser une analyse de régression avec un ordre plus haut, en utilisant une régression polynomiale de 2^{ème} et de 3^{ème} ordre, afin de mieux mettre en évidence ces trajectoires (e.g., la modélisation des trajectoires de F_2 de /u:/ dans l'environnement /d/ dans la situation de production Mot, voir Figure 45).

L'amélioration des taux de classification correcte des analyses discriminantes est visible au niveau des trois systèmes. Notre hypothèse de départ expliquait que c'est en arabe que la **dynamique** pourrait faciliter la perception des voyelles. Nos résultats ont montré qu'en plus de l'AM et de l'AJ, la caractérisation **dynamique** des voyelles en FR est pertinente au niveau de la classification de ses voyelles. Ces résultats vont dans le sens des propositions de René Carré (voir Carré, 2004, 2007 (à paraître) ; Carré *et al.*, 2004) où toute voyelle peut être caractérisé par sa dynamique spectrale, i.e., des trajectoires formantiques (voir, p. 76).

4.6. CONCLUSIONS

Nous venons d'examiner les résultats obtenus en production des voyelles de l'AJ, de l'AM et du FR. Une description approfondie des systèmes vocaliques des langues en **statique** (i.e., en n'évaluant que la valeur obtenue au milieu temporel de la voyelle, ou la « cible vocalique ») et en **dynamique** (i.e., en évaluant plusieurs valeurs obtenues à partir de l'**onset** vocalique jusqu'aux valeurs extraites au milieu temporel, par le calcul des **pentés formantiques**) a été proposée. Nous avons montré que l'utilisation des paramètres **dynamique** a permis de mieux caractériser les voyelles dans les trois langues, avec des taux moyens, de classification correcte par analyses discriminantes, améliorés de 10% à 30%. Ceci montre bien le rôle important de ces indices dans la caractérisation et la description des voyelles des trois langues.

Les résultats obtenus en **statique** et en **dynamique** dans les trois langues nous ont permis d'évaluer leurs systèmes vocaliques. Des concordances et des contradictions avec les propositions de la littérature ont été obtenues, à partir de nos échantillons de parole.

En AM par exemple, nos résultats ont permis de proposer un système vocalique composé des cinq voyelles /i: a: ə u u:/, où les deux voyelles brèves /ə u/ ont des statuts particuliers : quelle que soit la situation de production et dans les environnements /d d^ʕ/, ces deux voyelles ont fusionnées, tandis qu'elles sont réalisées comme deux voyelles distinctes dans les environnements /b k/. Des différences inter-individuelles ont été observées pour ce qui est de la réalisation de ces deux voyelles.

En AJ, nos résultats nous ont permis de proposer son système vocalique comme ayant les huit voyelles /i: e e: a: a o o: u:/, au lieu de */i: i e: a: a u o: u:/. L'originalité de ce système est le fait que les deux voyelles /i u/ ont été transformées et sont désormais produites par les locuteurs jordaniens comme /e o⁶⁷/, respectivement, quels que soient les lieux d'articulation, les situations de production et les locuteurs. La voyelle brève /a/ a été réalisée comme une voyelle [ɐ] dans la situation de production Mot seulement et quel que soit le lieu

⁶⁷ La réalisation de la voyelle /i/ comme [e] a été observée dans les trois situations de production et dans les environnements /b d^ʕ k/, dans l'environnement /d/, /i/ a été réalisée comme [i] à cause de l'item utilisé [dijja].

d'articulation, en revanche elle a été réalisée comme [a] dans les situations Syllabe et Isolation. Par conséquent, nous avons conclu que la structure syllabique CVC pouvait influencer sa réalisation (voir § 4.4.2.4 pour plus détails).

En FR, la littérature propose un système vocalique à onze voyelles orales : /i e ε a ɔ o u y ø œ/, or nos résultats nous permettent de proposer un système à dix voyelle /i e ε a ɔ o u y ø œ/. Pour ce qui est des deux voyelles ouvertes /a/ ne sont plus en opposition qualitative et forment désormais une seule et unique voyelle ouverte /a/, quels que soient le lieu d'articulation, la situation de production, le locuteur et le trait utilisé (i.e., **statique** ou **dynamique**). En ce qui concerne les voyelles « médianes » /e ε ɔ o ø œ/, nos résultats ont montré que ces voyelles se distinguent entre elles en **statique** et en **dynamique** quels que soient le lieu d'articulation, la situation de production, le locuteur et le trait utilisé (i.e., **statique** ou **dynamique**). Nous n'avons pas pu examiner les différences régionales en France, qui affectent directement le statut de ces voyelles entre le région Centre et Sud. Des études complémentaires peuvent mettre en évidence ces différences.

Nous voulons maintenant savoir jusqu'à quel point les auditeurs des trois langues sont sensibles aux différences observées entre les traits **statiques** et **dynamiques**. C'est le sujet du Chapitre 5 dans lequel nous examinons des résultats en perception des voyelles en AM, en AJ et en FR. Nous présentons les résultats de deux expériences de perception de la parole : une expérience de Recherche des "meilleures" instances – **prototypes** – vocaliques et une expérience de Catégorisation de l'espace vocalique. À chaque fois, des stimuli **statiques** et **dynamiques** sont proposés aux auditeurs afin d'en évaluer la pertinence. Nous terminons enfin par la présentation des résultats des analyses discriminantes qui nous permettrons d'évaluer les choix des auditeurs des trois langues dans l'identification des voyelles de leurs systèmes et la pertinence des systèmes vocaliques que nous proposons.

Chapitre 5.

RÉSULTATS EN PERCEPTION DES VOYELLES

5.1. INTRODUCTION

Nous avons vu dans le chapitre précédent que les traits **dynamiques** – et plus spécifiquement les informations contenues dans les transitions formantiques – permettent une meilleure caractérisation des voyelles à partir d’analyses du signal. Les différentes analyses de régression que nous avons appliquées sur nos données nous ont permis de mettre en évidence les trajectoires formantiques **intrinsèques** (dans la situation de production Isolation) et **extrinsèques** (dans les situations Mot et Syllabe) aux voyelles. Nous avons pu en examiner les effets sur les taux moyens de classification correcte des analyses discriminantes lesquels ont été améliorés en moyenne de 10% à 30% (selon le cas). La durée vocalique semble être également un trait primordial pour caractériser les voyelles en AM, AJ et même en FR. Nous avons vu également que les résultats obtenus en production des voyelles ne sont pas conformes avec ceux de la littérature et plus particulièrement au niveau du nombre et de la qualité des timbres vocaliques dans les trois langues.

En effet, pour certains auteurs, le système vocalique de l’AM est composé de trois voyelles longues /i: a: u:/ et de deux voyelles brèves /ə u/ et pour d’autres, le système vocalique bref n’est composé que d’une seule voyelle centrale /ə/ (voir § 1.4.2 pour plus de détails). Nos résultats ont montré que les deux voyelles brèves s’opposent significativement dans les trois environnements consonantiques /b k/ et « fusionnent » dans les environnements /d d^h/. Les taux de classification de l’analyse discriminante ont permis de séparer ces deux voyelles brèves. À partir de notre corpus, nous avons proposé que ce système est composé de cinq voyelles différentes avec des ellipses très grandes pour les deux voyelles brèves /ə u/. Le système vocalique en AJ est plutôt proposé avec les huit voyelles /i: i e: a a: o: u u:/ ; nos résultats ont montré que ce système se compose de huit voyelles qui diffèrent au niveau de la

qualité des timbres : /i: e e: a a: o: o u:/. L'originalité du système est que les voyelles brèves /i u/ en AJ sont différentes qualitativement – et quantitativement – de leurs correspondantes longues, /a/ quant à elle, est différente qualitativement uniquement en situation de production Mot ; en Syllabe et isolation, nous retrouvons le même timbre. Enfin en FR, le système vocalique est proposé avec les onze voyelles orales suivantes : /i e ε a ɔ o u y ø œ/ ; nous avons pu montrer que l'opposition entre les deux voyelles ouvertes /a ɔ/ n'est plus pertinente car tous les locuteurs français produisent ces voyelles, dans les trois contextes de présentation et dans les trois contextes de présentation comme une voyelle /a/. Pour les voyelles dites « médianes » en FR, nous n'avons pas examiné leurs différences qualitatives en fonction de l'origine régionale des locuteurs : dans la région Centre, les six voyelles « médianes » sont distinguées les unes des autres ; dans la région Sud, en revanche, ces six voyelles ne sont plus opposées et sont ainsi réalisées comme des voyelles mi-fermées. Dans notre étude, nous avons pu évaluer l'impact des différences régionales et des stratégies individuelles sur la réalisation de ces voyelles « médianes ». Au vu de ces remarques, nous proposons le système vocalique en FR comme ayant les dix voyelles orales suivantes : /i e ε a ɔ o u y ø œ/, sans la prise en compte des variations régionales au niveau des voyelles « médianes ».

Afin de mieux comprendre le rôle des traits **dynamiques** d'une part, et de vérifier la validité des résultats de production au niveau des variations qualitatives des timbres vocaliques dans les trois langues, d'autre part, nous avons effectué des expériences en perception des voyelles, dont les résultats sont présentés dans ce chapitre. Le but de ces expériences est de valider en particulier les effets de la densité des systèmes vocaliques sur la dispersion des voyelles de chaque système et de vérifier le rôle des indices **dynamiques** potentiels dans la caractérisation des voyelles des trois langues. Comme différents auteurs l'ont rappelé, l'espace vocalique perceptif est homogène entre les différents locuteurs, car la perception de la parole est un phénomène *many-to-one* où différents stimuli acoustiques phonétiquement différents sont identifiés par les auditeurs natifs de chaque langue comme un seul et unique stimulus auditif (voir les travaux de Johnson, 2000 ; Meunier, 2005 ; etc...).

Ce chapitre sera composé de quatre parties. Nous présentons dans la première partie notre méthodologie ; dans la deuxième les résultats en perception des "meilleures" instances -

prototypes – vocaliques ; dans la troisième les résultats des tests de catégorisation ; et la dernière établira une comparaison entre les résultats de production et ceux de perception des voyelles. Dans chaque partie, nous présentons les résultats dans les deux modalités **statique** et **dynamique** séparément, résultats qui seront comparés pour chacune des langues et entre celles-ci.

5.2. MÉTHODOLOGIE DU TRAVAIL

5.2.1. *Locuteurs*

Trente locuteurs masculins – différents de ceux de la phase de production – (10 locuteurs par langue) ont participé aux tests de perception des voyelles. Les locuteurs jordaniens sont originaires de la région d'Irbid au Nord de la Jordanie, les marocains sont de la ville de Fès et les français, de Lyon (ou de la Région Rhône-Alpes). Aucun trouble du langage, ni articulo-phonatoire ni auditif n'a été décelé chez les locuteurs qui présentent tous un audiogramme normal. L'intervalle d'âge est situé entre 18 et 30 ans. La situation socioprofessionnelle de tous les locuteurs est homogène, à savoir des étudiants de l'Université, issus d'un même milieu socioprofessionnel et vivant tous en ville. Tous les locuteurs sont soit entièrement monolingues (le cas des jordaniens), soit ils ont de faibles connaissances d'autres langues.

5.2.2. *Corpus*

La tâche des auditeurs des trois langues consiste à identifier les **prototypes** vocaliques de leurs systèmes. Au lieu de leur demander de choisir par exemple la meilleure voyelle /i/ ou /a/ isolément, nous avons choisi de leur faciliter la tâche en demandant de trouver la meilleure voyelle /i/ ou /a/ comme produite dans un mot précis. C'est pourquoi, nous avons élaboré trois corpus sensiblement proches de ceux utilisés dans la phase de production avec des modifications mineures (sauf en AM où un nouveau corpus a été élaboré, voir *infra*). Les structures syllabiques suivantes ont été utilisées à travers les trois langues : CV, CVC, CVCC, CVCV, CVCVC, CVCCV, CVCCVC ou VCVC. Les trois consonnes phonologiquement

communes entre les trois langues /b d k/ ont été utilisées⁶⁸ (voir les trois tableaux qui suivent).

Tableau 31 : Liste d'items utilisés pour la tâche de perception en AJ.

	b	sens	d	sens	k	sens
i:	bi:r	puits	di:nak	ta religion	ki:s	sac
i	bint	fille	dijja	indemnité	kilma	mot
e:	be:t	maison	de:r	monastère	ke:f	comment ?
a:	ba:b	porte	da:r	maison	ka:f	lettre "k"
a	bard	froid	dam	sang	kaf	paume
o:	bo:se	bisou	do:ri:	mon tour	ko:n	univers
u:	bu:ma	chouette	du:d	vers (de terre)	ku:sa:	courgette
u	burj	pont	dub	ours	kul	chacun

Tableau 32 : Liste d'items utilisés pour la tâche de perception en AM.

	b	sens	d	sens	k	sens
i	bmt	file	dirhim	dirham	kɪlma	mot
a	bard	froid	dam	sang	kalb	chien
o	burʒ	pont	dul	humiliation	kul	chacun
i:	bi:r	puits	di:b	loup	ki:f	comment ?
a:	ba:t	resté	da:b	fondre	ka:f	suffisant
u:	bu:l	pisse	du:d	vers (de terre)	ku:l	mange !!

Tableau 33 : Liste d'items utilisés pour la tâche de perception en FR.

	b	Transcription	d	Transcription	k	Transcription
i	bilan	bilā	dis	dī	qui	kī
e	bécasse	bekas	dé	de	képi	kepi
ɛ	baie	bɛ	dès	dɛ	caisson	kesō
a	bal	bal	date	dat	cale	kal
ɑ	basse	bas	d'âme	dām	case	kaz
ɔ	bol	bəl	dors	dəɞ	cote	kət
o	Baule	bol	dôme	dom	côte	kot
u	bout	bu	doux	du	coup	ku
œ	bœuf	bœf	odeur	ɔdœɞ	cœur	kœɞ
ø	bœufs	bø	vendeuse	vādøz	aqueuse	akøz
y	bu	by	du	dy	cul	ky

La nouvelle liste de mots proposés en AM a pour but de tester le statut phonologique des anciennes voyelles brèves héritées de l'arabe classique dans la variété actuelle de l'AM.

⁶⁸ Nous avons écarté le contexte /d^ɕ/ des expériences de perception, à cause de la difficulté de synthèse de ce contexte et des résultats de production que nous avons obtenus, concordant avec ceux de la littérature (dans ce travail et dans Al-Tamimi, 2002, Al-Tamimi & Barkat-Defradas, 2002 ; Barkat-Defradas *et al.*, 2003).

Comme nous l'avons expliqué auparavant, le système vocalique de l'AM est constitué de trois voyelles longues /i: a: u:/ et deux voyelles brèves /ə ʊ/. Les deux voyelles brèves de l'arabe classique /i a/ ont fusionné pour ne former qu'une seule voyelle centrale /ə/. Nous avons observé dans nos résultats en production que les locuteurs marocains avaient tendance à produire la voyelle /ə/ avec une aire de dispersion très grande allant de la position de la voyelle /ɪ/ à la voyelle /ɐ/, quant à la réalisation de la voyelle /ʊ/, elle est maintenue dans certains contextes consonantiques et chez certains locuteurs. Si, du point de vue de certaines études phonologiques, les deux voyelles /i/ et /a/ ont fusionné pour n'en former qu'une seule /ə/ et si l'opposition avec la voyelle /ʊ/ n'existe plus, il est très probable qu'au niveau phonétique ces voyelles soient toujours présentes chez certains locuteurs. Ceci peut être explicable soit par l'influence de la structure syllabique et/ou du contexte de présentation des consonnes adjacentes sur les voyelles, soit par des stratégies individuelles. C'est pourquoi nous allons tester leur présence en AM, en perception.

Pour ce qui concerne l'AJ, nous vérifierons si les auditeurs jordaniens sont sensibles aux variations qualitatives et quantitatives des timbres vocaliques de leur système : est-ce que les voyelles brèves de l'AJ /i a u/ sont identifiées par les auditeurs jordaniens comme ayant la même « position » dans l'espace vocalique que leurs correspondantes longues ; ou est-ce qu'elles sont identifiées avec une différence qualitative ? Enfin, pour le FR, nous testerons l'opposition entre les voyelles « médianes » d'un côté, et les deux voyelles ouvertes de l'autre. Le but de cette tâche de perception est d'étudier les connaissances qu'ont les auditeurs natifs et naïfs de leurs systèmes vocaliques.

5.2.3. Protocoles expérimentaux en perception

Deux types d'expériences en perception des voyelles ont été proposés : une expérience de Recherche de **prototypes** vocaliques et une expérience de Catégorisation des stimuli vocaliques dans l'espace vocalique de chaque langue. Pour les deux expériences, des traits **statiques** et **dynamiques** ont été utilisés.

Nous avons demandé aux auditeurs d'évaluer les deux expériences et les deux types de stimuli proposés, en leur demandant si les traits **dynamiques** ont facilité la caractérisation des **prototypes** et la catégorisation des stimuli dans l'espace vocalique de leur langue, et en leur demandant d'évaluer la difficulté des expériences effectuées. Étant donné que l'ordre de passation des expériences et l'effet d'entraînement et d'habituation peuvent affecter les choix des auditeurs. Pour contrebalancer ces effets d'ordre, nous avons permuté la présentation des expériences selon l'ordre suivant :

- * passation des expériences avec les stimuli **statiques** puis avec les stimuli **dynamiques** pour les auditeurs 1 à 5, mais avec les stimuli **dynamiques** puis les stimuli **statiques** pour les auditeurs 6 à 10,
- * passation des expériences de Recherche de **prototypes** vocaliques en premier et Catégorisation des espaces vocaliques ensuite pour les auditeurs 1, 3, 5, 6, 8 et 10 et l'inverse pour les auditeurs 2, 4, 7 et 9.

Nous présentons dans les parties qui suivent les protocoles expérimentaux des deux expériences de perception des voyelles.

5.2.3.1. Recherche de prototypes vocaliques

Ce test perceptif a pour rôle l'élaboration de l'espace acoustico-perceptif des locuteurs de chaque langue. C'est un test basé sur la méthode utilisée par Johnson *et al.* (1993), à savoir la méthode d'Ajustement de Formants, MOA (*Method Of Adjustment*). L'espace acoustique $F_1 \sim F_2$ est un espace synthétique continu et permet donc toutes les « synthèses » possibles à partir des valeurs de F_1 et de F_2 . Les valeurs de F_3 sont fixes et ont été déterminées sur la base des valeurs formantiques moyennes de voyelles produites par dix locuteurs par langue et correspondant aux valeurs de l'état stable proposées dans différents travaux (voir par exemple les travaux de Scholes, 1967 ; Hombert & Puech, 1984). La valeur de F_4 est restée constante à 3800 Hz pour toutes les voyelles. Les valeurs des quatre bandes passantes correspondant à celles des quatre premiers formants ont été déterminées chaque voyelle en s'inspirant des données de la littérature (Klatt & Klatt, 1990), (voir Tableau 34).

Tableau 34 : Valeurs de référence des quatre premiers formants (F₁, F₂, F₃ et F₄) et des quatre bandes passantes (B₁, B₂, B₃ et B₄) pour les voyelles en FR. Pour les voyelles /i: a: u:/ en arabe, les mêmes valeurs que pour leurs correspondantes en FR ont été proposées (/i ɐ ʊ/ = voyelles brèves en arabe).

voyelle	F ₁	F ₂	F ₃	F ₄	B ₁	B ₂	B ₃	B ₄
i	250	2500	3100	3800	50	100	210	200
ɪ	375	2250	2900	3800	50	100	210	200
e	350	2300	2900	3800	50	91	219	200
ɛ	500	2200	2800	3800	50	95	247	200
ɐ	600	1350	2900	3800	50	100	266	422
a	700	1600	2700	3800	50	100	266	422
ɑ	700	1200	2800	3800	50	100	218	332
ɔ	500	950	2400	3800	50	50	100	200
o	350	850	2300	3800	50	50	100	100
ʊ	375	1000	2300	3800	50	50	50	200
u	250	650	2200	3800	50	50	50	200
y	250	1900	2300	3800	50	100	266	422
ø	350	1600	2300	3800	50	50	100	200
œ	500	1600	2300	3800	50	100	200	200

Un programme informatique spécifique, développé en C++ en unicode et conçu par René Carré a été élaboré pour cette tâche. Avec l'interface du programme (voir Figure 113), l'auditeur dispose d'une référence pour chaque voyelle examinée (e.g., pour une voyelle /i/, les valeurs formantiques de production de la voyelle en isolation ont été proposées pour donner à l'auditeur des repères dans cet espace continu et ainsi lui simplifier la tâche)⁶⁹. Nous avons choisi de présenter des valeurs des quatre premiers formants et des quatre premières bandes passantes communes dans les trois langues : nous cherchons ainsi à identifier le rôle possible de la densité des systèmes sur la taille des espaces vocaliques dans chaque langue, d'une part, et à évaluer le rôle du changement de contexte de présentation des consonnes adjacentes sur la dispersion des voyelles, de l'autre.

⁶⁹ Nous nous sommes rendu compte que, sans repère dans un espace bidimensionnel, les auditeurs passaient beaucoup du temps à chercher à se familiariser avec le test, et perdaient ainsi patience dans les passations des expériences. Chaque stimulus entendu les obligeait à se familiariser à nouveau avec l'interface du programme pour chercher des repères.

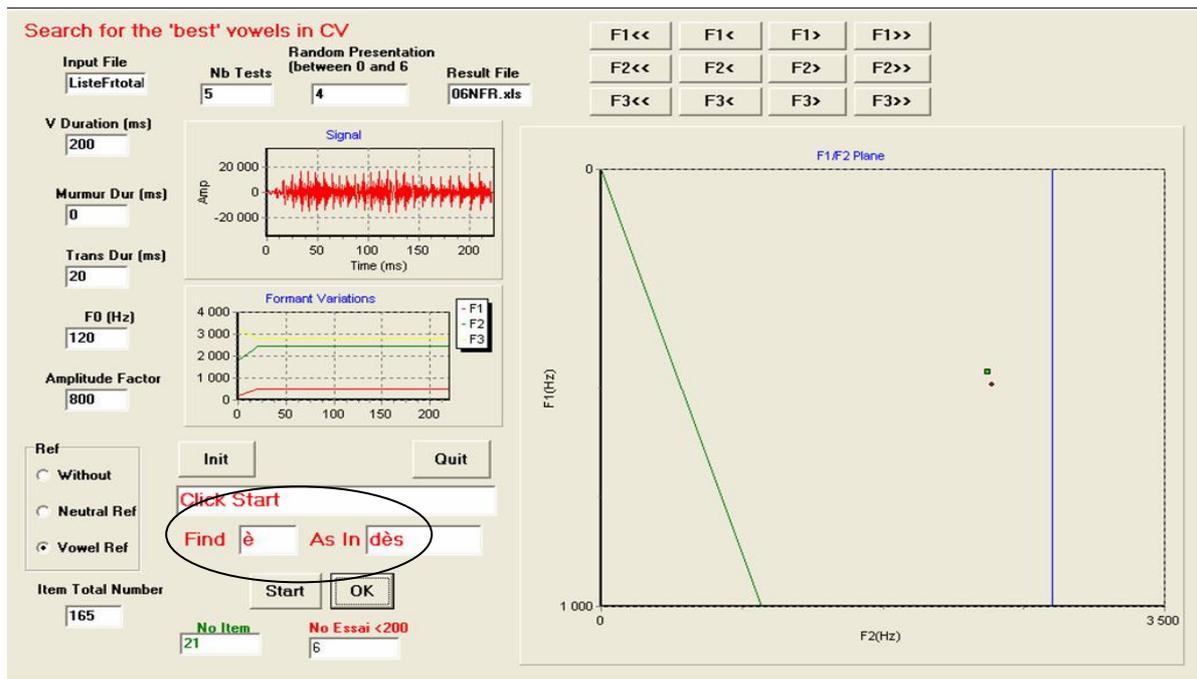


Figure 113 : Interface du programme de perception des voyelles en **dynamique**, présentant les différents items.

L'auditeur se déplace dans l'espace acoustique synthétique pour retrouver la meilleure représentation du **prototype** vocalique associée à la voyelle présente dans le mot (partie marquée par une ellipse). Dans l'espace acoustique de F_1 - F_2 , le point avec la forme en losange indique le dernier choix effectué par l'auditeur ; celui en carré, le précédent choix du même stimulus.

Nous avons proposé une phase de familiarisation et d'entraînement pour chaque auditeur, où une liste de cinq mots différents a été proposée. Une fois cette phase terminée, l'auditeur a été invité à manipuler l'interface du logiciel qui lui permettait, d'un simple click de souris, d'écouter différents stimuli synthétiques représentant les voyelles avec des traits **statiques** et **dynamiques**. Les traits **statiques** (voir Tableau 34) utilisés correspondent aux valeurs cibles de l'état stable, avec une durée vocalique constante de 200 ms (200 ms pour les voyelles longues et 100 ms pour les voyelles brèves en arabe) et un f_0 montant~descendant de 120-132-96 Hz (valeurs correspondant à l'**onset**, au milieu temporel et à l'**offset** de chaque voyelle). Comme pour la phase de production, nous avons adapté l'interface du programme pour les auditeurs arabophones : nous avons affiché la voyelle et le mot en caractères arabes, ainsi l'examineur n'avait aucune influence directe sur l'auditeur.

En ce qui concerne les traits **dynamiques**, ils sont la résultante de l'association à la fois d'une transition formantique précise et d'un stimulus représenté par les traits **statiques** cités plus haut. Nous nous sommes appuyé sur les données de la littérature concernant le « locus virtuel » (Delattre *et al.*, 1955, voir § 4.3.2.1.1) et sur des tests perceptifs, à partir de

différents programmes développés par René Carré⁷⁰, afin de choisir les valeurs du meilleur « locus » bilabial, dental et vélaire pour un locuteur donné pour les quatre premiers formants (voir Tableau 35). Ainsi, les consonnes synthétisées correspondent aux consonnes initiales présentes dans les mots écrits sur l’interface du logiciel. Une fois les valeurs de transitions sélectionnées, nous avons soumis les stimuli élaborés à des tests de jugement de qualité des consonnes initiales, effectués par nous-même et par un auditeur natif de chaque langue.

Tableau 35 : Valeurs de référence du « locus » vers lequel les transitions des quatre premiers formants convergent, en contexte bilabial, dental et vélaire pour toutes les voyelles présentes dans les trois langues.

voyelle	Bilabial				Dental				Vélaire			
	F ₁	F ₂	F ₃	F ₄	F ₁	F ₂	F ₃	F ₄	F ₁	F ₂	F ₃	F ₄
i	175	1200	2000	2500	175	1800	3200	3900	175	2800	2950	3600
ɪ	175	1200	2000	2500	175	1800	3200	3900	175	2300	2700	3600
e	175	1200	2000	2500	175	1800	3200	3900	175	2600	2900	3600
ɛ	175	1200	2000	2500	175	1800	3200	3900	175	1400	2800	3600
ɐ	350	1200	2000	2500	350	1800	3200	3900	350	2200	2500	3600
a	350	1200	2000	2500	350	1800	3200	3900	350	2200	2500	3600
ɑ	350	1100	1900	2500	350	1800	3200	3900	350	2100	2400	3600
ɔ	175	1100	1900	2500	175	1800	3200	3900	175	1300	1900	3600
o	175	1100	1900	2500	175	1800	3200	3900	175	1200	1900	3600
ɒ	175	1300	1900	2500	175	1800	3200	3900	175	900	1900	3600
u	175	1100	1900	2500	175	1800	3200	3900	175	1000	1900	3600
y	175	1200	1800	2500	175	1800	3200	3900	175	2000	2050	3600
ø	175	1200	1800	2500	175	1800	3200	3900	175	1900	2000	3600
œ	175	1200	2000	2500	175	1800	3200	3900	175	1900	2000	3600

La présentation des stimuli a été effectuée aléatoirement avec cinq répétitions. Chaque auditeur avait la possibilité d’écouter jusqu’à 200 stimuli « différents » pour chaque **prototype** vocalique qu’il devait identifier avant de valider son choix : ces différents stimuli variaient de quelques Hertz sur le plan F₁ et/ou F₂ en fonction des déplacements opérés par

⁷⁰ Nous avons utilisé deux programmes différents, conçus par René Carré et basés sur le synthétiseur à formants développé par Klatt & Klatt (1990), qui nous ont permis de trouver les « meilleurs Locus virtuels » pour les F₂ et F₃ : PercepCV et PlagePercepCV. Le principe de base de ces deux programmes est d’effectuer une catégorisation du contexte de présentation des consonnes perçues dans une syllabe CV en partant d’une voyelle précise. Les résultats sont enregistrés dans un fichier excel.

l'auditeur. Nous avons choisi d'enregistrer ces choix pour chaque auditeur et pour chaque stimulus, afin d'évaluer, les directions de déplacement de l'auditeur pour choisir les **prototypes**, d'une part, et d'identifier les voyelles les plus difficiles à choisir, de l'autre.

Pour ce qui est des stimuli **statiques**, seule la présentation visuelle a été modifiée (i.e., l'auditeur entend trois fois le même stimulus pour la voyelle /i/ en FR par exemple, pour trois mots différents : bilan, dit, qui) : dans ce cas, nous parlerons d'effet du contexte de présentation. En revanche, au niveau des stimuli **dynamiques**, l'auditeur entendait trois stimuli différents pour les trois mots : la différence était représentée par la partie initiale de chaque stimulus et la voyelle était identique pour les trois contextes de présentation (i.e., mêmes fréquences). L'objectif poursuivi en utilisant un corpus écrit et demandant aux auditeurs de choisir les **prototypes** vocaliques de leurs systèmes « comme produits dans » le mot associé était de chercher à savoir si le choix des voyelles **prototypiques** allait être influencé par la possible « représentation mentale » de l'environnement consonantique constitué par un mot écrit ou si les auditeurs allaient choisir des voyelles **prototypiques** « invariantes » (i.e., sans la prise en compte des influences consonantiques).

5.2.3.2. Catégorisation de l'espace vocalique

Ce test permet de trouver les limites ou les frontières entre les catégories vocaliques de la langue étudiée. Nous avons effectué la synthèse d'un jeu de 56 stimuli, avec une durée de 200 ms (et de 100 ms pour les voyelles brèves de l'arabe) et un f_0 montant~descendant de 120-132-96 Hz, grâce à un logiciel de synthèse à formants élaboré par René Carré (s'appuyant sur le synthétiseur à formants de Klatt & Klatt, 1990)⁷¹. L'espace vocalique synthétique est constitué des combinaisons suivantes des deux premiers formants : F_1 varie de 250 à 750 Hz

⁷¹ Ce programme : SyntFormVoy, conçu par René Carré, nous a permis de synthétiser tous les stimuli **statiques** et **dynamiques**. Ce programme permet de synthétiser une voyelle et jusqu'à quinze voyelles différentes en modifiant la durée de transition entre les voyelles et la durée vocalique. Nous l'avons adapté à nos stimuli **dynamiques** en modifiant les valeurs formantiques des quatre premiers formants, des quatre bandes passantes, des durées de transition et des durées de la « voyelle » en remplaçant par la durée de la transition : ainsi pour la première partie du stimulus, nous avons choisi 0 comme durée de transition et 25 ms comme durée de « voyelle ». Dans notre étude, toutes les consonnes initiales ont été synthétisées à partir des transitions seulement, i.e., sans « burst ».

avec un pas de 100 Hz, et F_2 varie de 650 à 2350 avec un pas de 200 Hz (voir Tableau 36). Une série de voyelles arrondies et non-arrondies (stimuli N°. 47 à 56) pour les voyelles antérieures et postérieures a été synthétisée en variant F_3 (i.e., avec deux valeurs à 2300 Hz et 2700 Hz). Cette expérience est une réplique de celle établie par Hombert & Puech (1984).

Tableau 36 : Valeurs des quatre premiers formants des 56 stimuli synthétiques utilisées dans l'expérience de Catégorisation de l'espace vocalique.

N°. Stimulus	F ₁	F ₂	F ₃	F ₄	N°. Stimulus	F ₁	F ₂	F ₃	F ₄
1	250	2350	3100	3800	29	550	1750	2900	3800
2	250	2150	3100	3800	30	550	1500	2500	3800
3	250	1950	2900	3800	31	550	1250	2500	3800
4	250	1750	2900	3800	32	550	1050	2300	3800
5	250	1500	2500	3800	33	550	850	2300	3800
6	250	1250	2500	3800	34	550	650	2200	3800
7	250	1050	2300	3800	35	650	1950	2900	3800
8	250	850	2300	3800	36	650	1750	2900	3800
9	250	650	2200	3800	37	650	1500	2500	3800
10	350	2350	3100	3800	38	650	1250	2500	3800
11	350	2150	3100	3800	39	650	1050	2300	3800
12	350	1950	2900	3800	40	650	850	2300	3800
13	350	1750	2900	3800	41	750	1950	2900	3800
14	350	1500	2500	3800	42	750	1750	2900	3800
15	350	1250	2500	3800	43	750	1500	2500	3800
16	350	1050	2300	3800	44	750	1250	2500	3800
17	350	850	2300	3800	45	750	1050	2300	3800
18	350	650	2200	3800	46	750	850	2300	3800
19	450	2150	3100	3800	47	250	1950	2300	3800
20	450	1950	2900	3800	48	350	1950	2300	3800
21	450	1750	2900	3800	49	450	1750	2300	3800
22	450	1500	2500	3800	50	550	1750	2300	3800
23	450	1250	2500	3800	51	650	1500	2300	3800
24	450	1050	2300	3800	52	750	1500	2300	3800
25	450	850	2300	3800	53	250	850	2700	3800
26	450	650	2200	3800	54	350	850	2700	3800
27	550	2150	3100	3800	55	450	850	2700	3800
28	550	1950	2900	3800	56	550	850	2700	3800

Deux séries de stimuli ont été proposées aux auditeurs : des stimuli **statiques** et des stimuli **dynamiques** (avec une transition formantique de 25 ms, adaptée au seul contexte dental et ayant les valeurs suivantes du « locus virtuel » des quatre premières transitions formantiques : 175, 2000, 3200 et 4000 Hz)⁷².

⁷² Nous avons choisi un seul contexte consonantique pour réduire la durée totale de passation des expériences perceptives (Recherche de **prototypes** vocaliques et Catégorisation des espaces vocaliques, en utilisant les indices **statiques** et **dynamiques** à chaque fois).

Un programme informatique spécifique, développé en Tcl/Tk (version 8.4.3.0, 2003) et Snack Sound ToolKit (version 2.2 ; Sjölander, 2002) et conçu par François Pellegrino, a été élaboré pour cette tâche. L'interface du programme permet à l'auditeur de lire les consignes et de s'identifier (en utilisant le même code proposé dans les autres expériences : un code à 5 caractères composés du code de la langue (AM, AJ ou FR), du symbole M pour masculin et du numéro de l'auditeur en deux chiffres).

La tâche des auditeurs consistait à identifier chaque stimulus entendu et à le « placer » dans une catégorie vocalique représentée par un des mots de l'environnement /d/ de la liste d'items présentés dans les tableaux 40 à 42 (e.g., l'auditeur entend un stimulus ayant les fréquences suivantes : 250, 2350, 3100 et 3800 pour F1, F2, F3, et F4, respectivement ; il l'identifie comme une voyelle [i] et le place dans la catégorie « dis » en FR). L'auditeur a la possibilité de réécouter le stimulus autant de fois qu'il le souhaite. Une pause a été programmée tous les 50 stimuli et l'auditeur était invité à se déplacer, à quitter la salle, ou simplement à se reposer pour le temps qu'il le souhaitait. Si le stimulus n'était pas identifié par l'auditeur, et par conséquent non « placé » dans une catégorie vocalique quelconque, l'auditeur était invité à choisir la case indiquant « aucun ne correspond » (voir Figure 114). Cette expérience a été conçue pour permettre à l'auditeur d'être à 100% autonome, (i.e., sans aucune intervention extérieure). Ainsi, nous l'avons adapté à la langue arabe, où les consignes étaient rédigées entièrement en arabe (voir Figure 114, partie b). Nous avons choisi d'utiliser l'**arabe standard contemporain** pour l'arabisation de la tâche et pour le contact avec les locuteurs.

Test de reconnaissance n° 1

Dans quel mot le son entendu est-il présent ?

dis	dé
dés	date
d'âme	dors
dôme	doix
du	ventouse
odeur	
Aucun ne correspond	

Quitter

Écouter le test suivant

Ré-écouter le son

a)

Test de reconnaissance n° 1

الصوت الذي لم سمعته موجود في أي كلمة من القائمة الموجودة أمامك؟

دين	دير
ديه	دار
ديم	دور
ديب	دود
لصوت المسموع غير موجود في القائمة!	

Quitter

إستبدل هذا للصوت التالي و اإستمع للصوت التالي

إستمع للصوت مرة أخرى

b)

Figure 114 : Interfaces du programme de Catégorisation de l'espace vocalique en FR (partie a) et en AJ (partie b), avec la liste des mots qui représentent chacun une voyelle précise en contexte dental.

Nous présentons dans les parties qui suivent, les résultats obtenus en perception des voyelles en AM, en AJ et en FR. Nous procédons ensuite à une comparaison entre les trois systèmes en étudiant les trois voyelles communes /i a u/.

5.3. RECHERCHE DE PROTOTYPES VOCALIQUES

Nous présentons dans cette partie les résultats des expériences de Recherche de **prototypes** vocaliques synthétisés isolément (désormais nous utilisons le terme **statique** pour désigner cet environnement) ou en contexte consonantique de type CV (désormais nous utilisons le terme **dynamique** pour désigner cet environnement) Nous commençons par la présentation des analyses et du traitement des données, ensuite par les résultats pour l'AM, l'AJ et le FR et terminons par une comparaison inter-langues des trois voyelles /i a u/.

5.3.1. *Analyse et traitement des données*

Le corpus obtenu contient 7500 **prototypes** vocaliques qui ont été identifiés par les auditeurs des trois langues (1800 **prototypes** en AM, 2400 en AJ et 3300 en FR), en **statique** et en **dynamique**. Les valeurs formantiques de F_1 et F_2 obtenues ont été converties en Bark en suivant la méthode utilisée précédemment dans le traitement des données en production. Nous avons ensuite calculé les moyennes et les écart-types de ces valeurs par langue, par type de stimuli (**statique** et **dynamique**), par contexte de présentation (bilabial, dental et vélaire) et par voyelle.

Comme en production des voyelles, nous avons appliqué une analyse statistique : une MANOVA à cinq facteurs : langue (3), sujets (10), modalité (**statique** vs **dynamique**) (2), contexte de présentation (3) et voyelles (six en AM, huit en AJ et onze en FR). Le but de cette analyse était de vérifier la validité statistique de nos résultats. Ensuite nous avons utilisé une méthode de classification des données : une analyse discriminante, qui a pour but de tester la validité des résultats obtenus en fonction d'une variable précise, afin de valider les résultats (pour plus de détails sur ces deux méthodes d'analyses, voir Chapitre 4, § 4.3.1.1).

5.3.2. *En Arabe Marocain*

Nous divisons la présentation des résultats de Recherche de **prototypes** vocaliques en AM en quatre parties : les résultats en **statique**, en **dynamique**, issus de la comparaison entre les deux modalités et issus de l'analyse discriminante.

5.3.2.1. Résultats en statique (V Target)

Les résultats des **prototypes** vocaliques identifiés par les auditeurs marocains sont présentés sur la Figure 115. Les résultats graphiques ne montrent pas de différences significatives en fonction du contexte de présentation (présent uniquement sous forme visuelle) : sur F_1 : $F(2, 882)=1,61$; $p=0,2$; et sur F_2 : $F(2, 882)=0,6$; $p=0,55$. Les résultats de l'interaction entre les voyelles et le contexte de présentation montrent qu'aucun effet n'est obtenu sur les deux axes (sur F_1 : $(F(10, 882) = 1,24$; $p<0,26$; sur F_2 : $F(10, 882) = 1,01$; $p=0,43$). Les résultats du calcul de taille de l'aire de dispersion⁷³ confirment cette tendance (voir Tableau 37) : nous obtenons une moyenne non significative de 1 Bark² entre les contextes dental et bilabial ($F(1, 18)=0,55$; $p=0,47$), et de 1,25 Bark² entre le dental et le vélaire ($F(1, 18)=1,04$; $p=0,32$).

Ces résultats indiquent que les espaces vocaliques des **prototypes** identifiés par les auditeurs marocains ne sont pas significativement différents. Nous nous attendions à ne pas observer de différence significative entre les **prototypes** vocaliques choisis par les auditeurs marocains. L'explication de ce résultat est la ressemblance auditive entre les trois stimuli proposés : les auditeurs percevant des stimuli synthétiques identiques qui ne changent qu'au niveau du stimulus visuel ne choisissent pas de **prototypes** différents en fonction du contexte de présentation. Ce résultat va dans le sens des explications données dans la littérature sur la stabilité de l'espace vocalique perceptif : la perception de la parole est un phénomène *many-to-one* (voir les travaux de Meunier (2005), entres autres).

⁷³ Nous utiliserons désormais les termes : **aire de dispersion des voyelles** pour parler de la taille de l'espace vocalique obtenue à partir des voyelles étudiées ; et **ellipse de dispersion des voyelles** pour parler de la taille des ellipses de chaque voyelle.

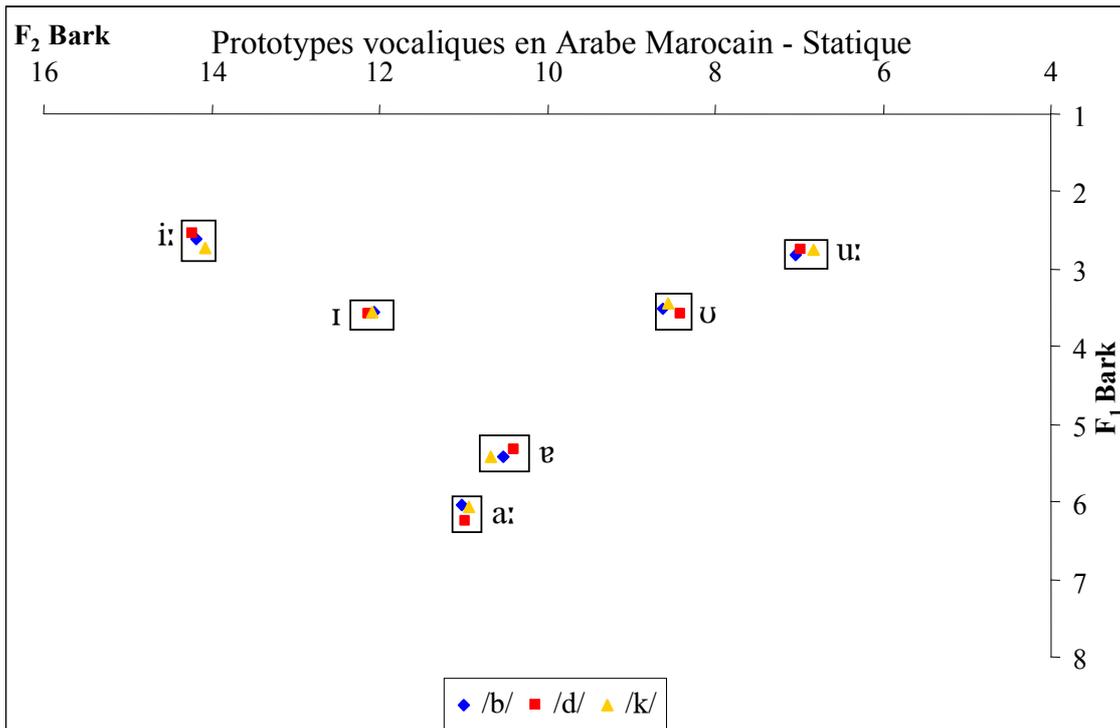


Figure 115 : Espace vocalique des **prototypes** vocaux identifiés en AM **statique**, en fonction des contextes de présentation.⁷⁴

Tableau 37 : Taille des aires de dispersion des voyelles **prototypiques** identifiées en AM en **statique**, en fonction des contextes de présentation (en Bark²).

	Moyenne	Écart-Type
Bilabial	17,08	2,63
Dental	18,01	3,00
Vélaire	16,75	2,53

La position des six voyelles **prototypiques** obtenue sur la Figure 115 montre des différences significatives liées au timbre des voyelles (i.e., des voyelles différentes sur le plan perceptif) et ce sur F₁ : $F(5, 882)= 1219,23$; $p<0,001$; et sur F₂ : $F(5, 882)=2255,43$; $p<0,001$. Les trois voyelles longues /i: a: u:/ sont positionnées aux extrémités de l'espace vocalique ; les trois voyelles brèves /I ɐ U/ obtiennent des positions différentes sur l'espace vocalique s'éloignant de celle d'une voyelle [ə]. Nous obtenons trois voyelles **prototypiques**

⁷⁴ Les carrés dessinés autour de chaque catégorie vocale sont affichés simplement par soucis de visibilité pour regrouper les données par timbre vocalique. Leur taille est arbitraire et donc sans lien avec la dispersion des ellipses.

différentes quel que soit le contexte de présentation, sans aucune « centralisation » dans l'espace acoustico-perceptif. Comme nous l'avons vu, la littérature propose que les deux voyelles brèves /i a/ en AM ont fusionné et forment une seule voyelle centrale /ə/ tandis que la voyelle brève /u/ est restée en opposition de longueur vocalique avec /u:/, avec une « centralisation » dans le plan $F_1\sim F_2$, et est devenue /ʊ/. Nos résultats montrent que les auditeurs marocains perçoivent trois timbres vocaliques brefs différents.

Nous observons néanmoins des différences significatives entre les choix des voyelles **prototypiques** en AM pour toutes les voyelles en fonction de l'auditeur, différences observables sur les deux axes : sur F_1 : $F(9, 882)=58,6$; $p<0,001$; et sur F_2 : $F(9, 882)=16,72$; $p<0,001$ (voir Figure 116). Afin de quantifier cette différence, nous avons calculé la taille des ellipses de chaque catégorie vocalique. Les résultats, présentés sur la Figure 117, montrent une variabilité inter-individuelle peu importante pour les deux voyelles longues /i: u:/ dans les trois contextes de présentation, mais une variabilité bien marquée pour les autres voyelles dans les trois contextes de présentation sauf pour la voyelle /ʊ/ qui affiche une dispersion très grande en contexte vélaire. En moyenne, ces ellipses sont plus grandes en contexte vélaire, plus petites en contexte dental et intermédiaires en contexte bilabial (différences non significatives). Nous obtenons une interaction significative entre les auditeurs et les voyelles sur les deux axes : sur F_1 : $F(45, 882) = 15,58$; $p<0,001$; sur F_2 : $F(45, 882) = 5,55$; $p<0,001$. Ceci peut être expliqué par des différences de stratégies individuelles : les auditeurs marocains organisent leurs espaces acoustico-perceptifs différemment.

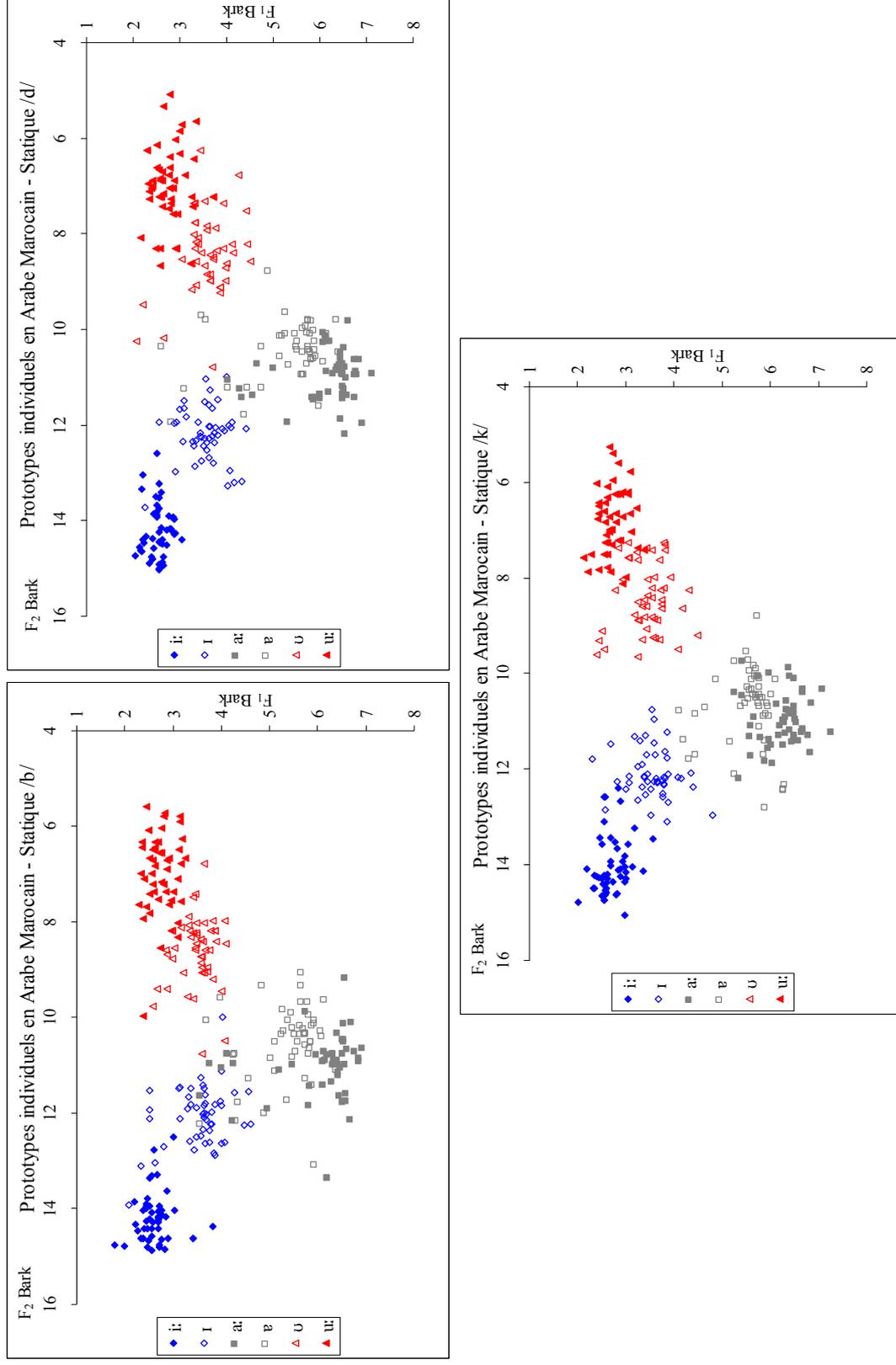


Figure 116 : Dispersion des voyelles **prototypiques** individuelles en AM en **statique**, en fonction des contextes de présentation.

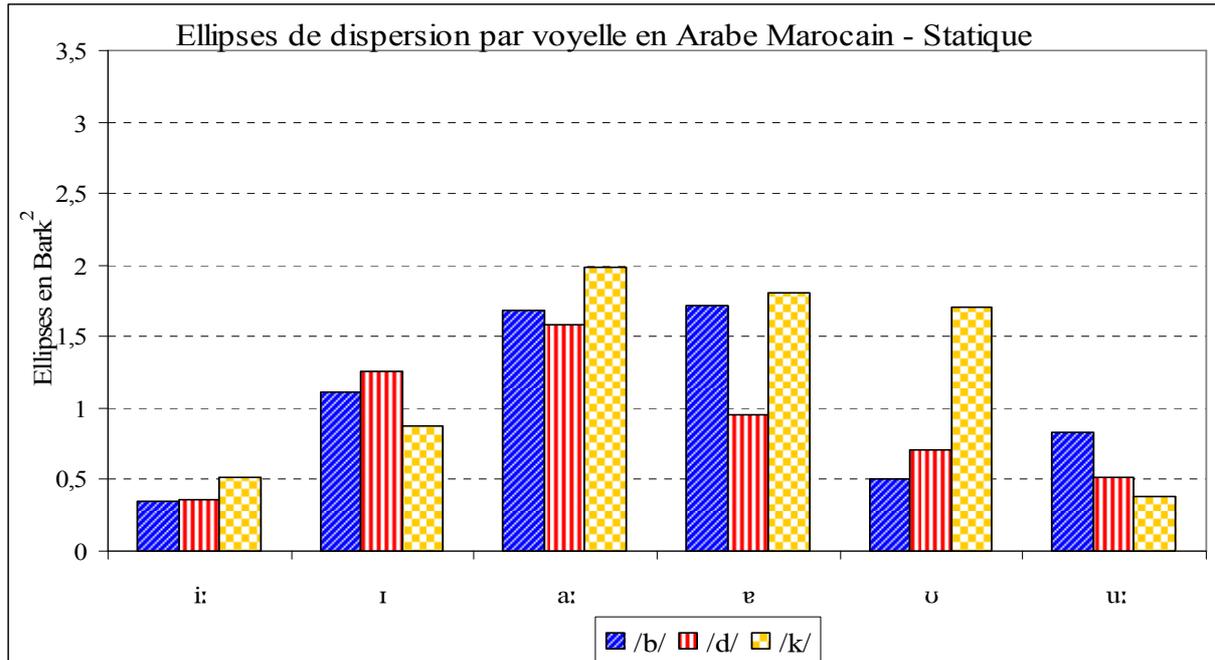


Figure 117 : Taille des ellipses de dispersion par voyelle **prototypique** perçue en AM en **statique**, en fonction des contextes de présentation.

La littérature explique que la taille des ellipses de dispersion de chaque voyelle et plus particulièrement pour les voyelles extrêmes /i: a: u:/, est fonction de la densité des systèmes vocaliques. Le système vocalique perceptif de l'AM atteste six voyelles différentes. Nous observons que seules les deux voyelles extrêmes /i: u:/ obtiennent les ellipses les plus petites, /a:/ en revanche obtient l'ellipse la plus grande. Ces résultats vont dans le sens des propositions de la **théorie quantique** (pour /i: u:/) et de la **théorie de la dispersion adaptée** (pour /a:/). Nous avons proposé aux auditeurs marocains des stimuli longs et brefs (200 ms et 100 ms, respectivement) afin d'être cohérent avec les données de la littérature concernant l'opposition de quantité en arabe. La différence entre les stimuli proposés était fonction de la durée variable et de changement de qualité phonétique des timbres. Tous les auditeurs marocains nous ont révélé ne pas avoir fait usage de l'opposition de durée de stimuli dans les expériences proposées pour choisir les **prototypes** vocaliques correspondants aux voyelles « longues » et « brèves ». Ils n'étaient pas conscients, d'après eux, de cette variation, car ils cherchaient avant tout à identifier le « timbre » correspondant à chaque stimulus. Dans tous les cas, le protocole proposé ne permettait pas aux auditeurs de modifier la durée de stimulus comme ils le souhaitaient lorsqu'une voyelle longue ou brève était à identifier.

Les résultats obtenus en **statique** montrent que les auditeurs marocains sont sensibles uniquement aux variations liées aux timbres vocaliques des six voyelles proposées ; aucune différence significative n'a été obtenue pour la variation liée au contexte de présentation (variation disponible sur le plan visuel uniquement), ni en fonction de la durée des stimuli. Nous voulons savoir comment les traits **dynamiques** vont aider ou pas les auditeurs marocains à identifier les voyelles de leur système.

5.3.2.2. Résultats en dynamique (CV Target)

Les résultats des choix des **prototypes** vocaliques identifiés par les auditeurs marocains en **dynamique** sont présentés graphiquement sur la Figure 118. Les espaces vocaliques obtenus sont significativement différents en fonction des contextes de présentation c'est dans l'environnement vélaire que l'espace est plus « ouvert » sur l'axe F_1 : $F(2, 882)=7,11$; $p<0,001$; tandis que c'est en contexte bilabial qu'il est plus « postérieures », sur l'axe F_2 : $F(2, 882)=25,21$; $p<0,001$.

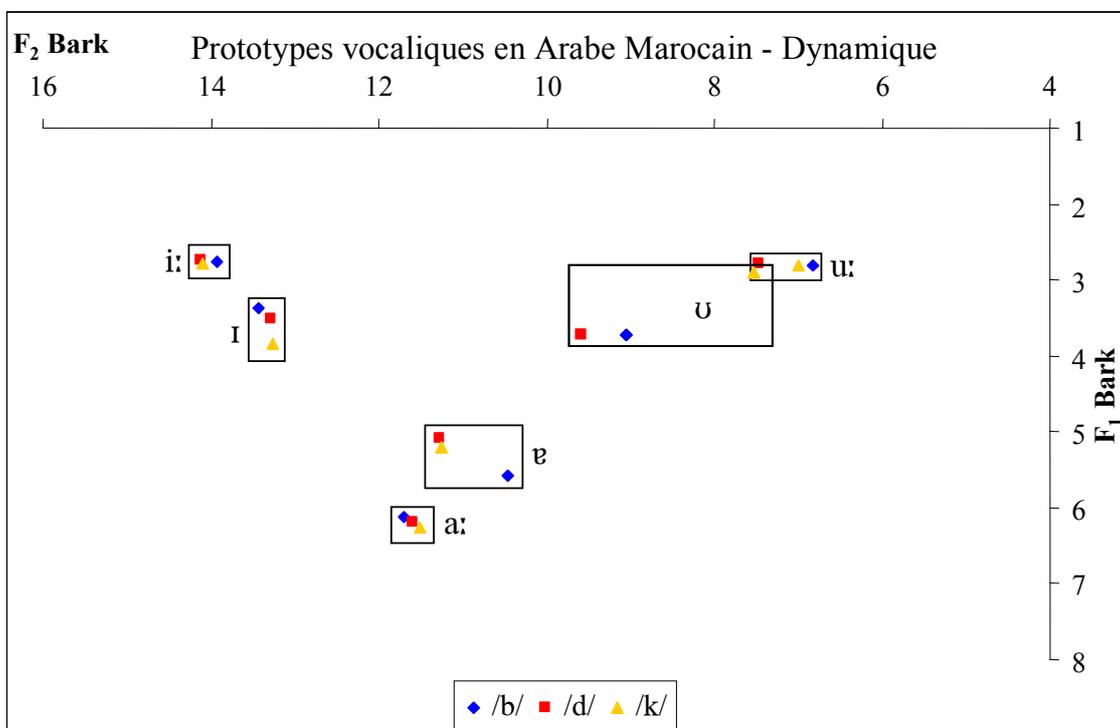


Figure 118 : Espace vocalique des **prototypes** vocaliques identifiés en AM en **dynamique**, en fonction des contextes de présentation.

Nous n'observons pratiquement aucune différence quant à la taille des aires de dispersion des voyelles **prototypiques** identifiées dans les trois contextes consonantiques (voir Tableau 38).

Tableau 38 : Taille des aires de dispersion des voyelles **prototypiques** identifiées en AM en **dynamique**, en fonction des contextes de présentation (en Bark²).

	Moyenne	Écart-Type
Bilabial	17,62	2,26
Dental	17,59	1,63
Vélaire	17,66	1,82

Nous obtenons des différences significatives liées aux timbres des voyelles, sur F_1 : $F(5, 882)=1594,25$; $p<0,001$; et sur F_2 : $F(5, 882)=2035,31$; $p<0,001$. Les résultats de l'interaction entre les voyelles et le contexte de présentation donnent des différences significatives : sur F_1 , les voyelles /i: ɪ a:/ identifiées en contexte vélaire sont légèrement plus « ouvertes » que dans les deux autres contextes ($F(10, 882) = 13,82$; $p<0,001$) ; sur l'axe F_2 , seules les voyelles /i: ɛ u:/ identifiées en contexte bilabial sont plus « postérieures » ($F(10, 882) = 18,86$; $p<0,001$, (voir Figure 118). Comme en **statique**, nous obtenons des différences significatives liées aux timbres des trois voyelles brèves en AM ; i.e., /i a/ ne constituent pas une seule et unique voyelle centrale /ə/ mais deux timbres distincts. Nous observons par contre un rapprochement très important entre les **prototypes** vocaliques de /ʊ/ et /u:/ en contexte vélaire : résultat obtenu également en production des voyelles (voir § 4.4.1).

Nous obtenons des différences significatives entre les auditeurs marocains dans les choix des **prototypes** vocalique de leur système visibles sur F_1 : $F(9, 882)= 16,94$; $p<0,001$; et sur F_2 : $F(9, 882)=16,01$; $p<0,001$ (voir Figure 119).

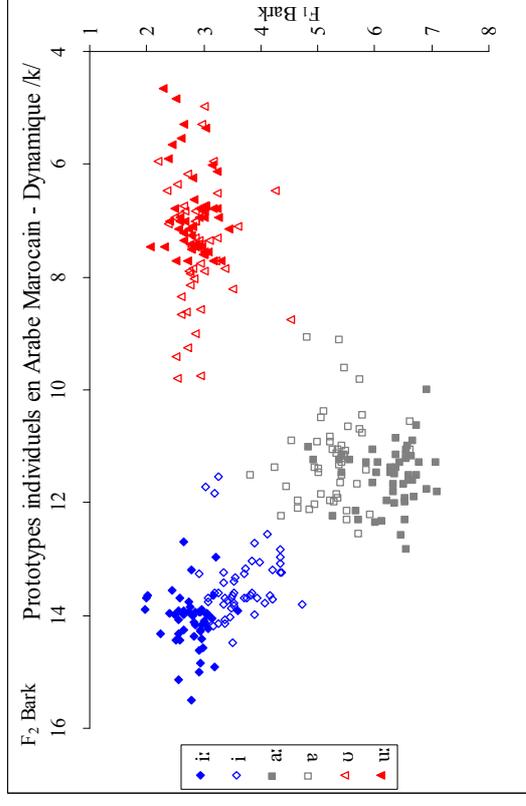
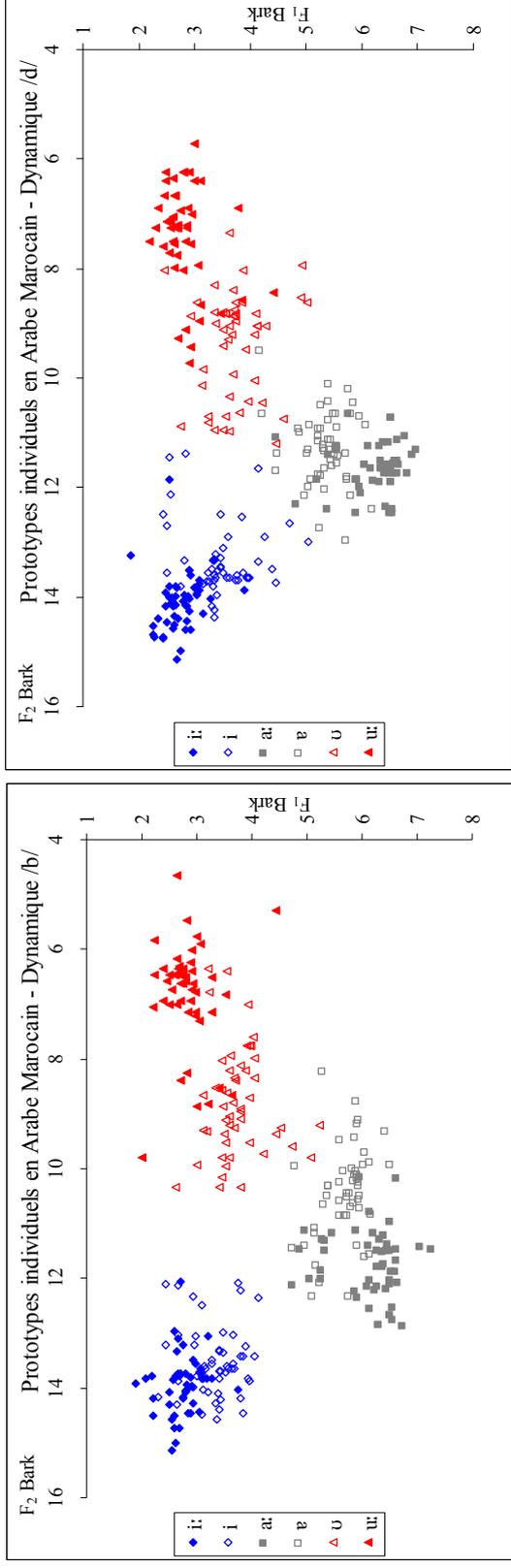


Figure 119 : Dispersion des voyelles **prototypiques** individuelles en AM en **dynamique**, en fonction des contextes de présentation.

À partir des résultats individuels, nous observons une homogénéité entre les voyelles **prototypiques** longues choisies par les auditeurs marocains. Les **prototypes** vocaliques brefs quant à eux, sont identifiés avec plus de variation. En effet, la voyelle **prototypique** /ɪ/ est tantôt perçue comme une voyelle [i] (en contexte /b/), tantôt comme une voyelle allant des voyelles [i] à [e] (en contextes /d/ et /k/). La voyelle /ɘ/ est identifiée soit comme une voyelle « postérieure » [ɑ] ou [ɔ] (en contexte /b/), soit comme une /a/ centrale (en contextes /d/ et /k/). Enfin, la voyelle /ʊ/ est identifiée comme une voyelle « centrale » [ə] (en contextes /b/ et /d/) et comme une voyelle soit « postérieure » /ʊ/, soit comme une voyelle [ə] plus « fermée » (en contexte /k/). Nous obtenons une interaction significative entre les auditeurs et les voyelles sur les deux axes : sur F_1 : $F(45, 882) = 6,91$; $p < 0,001$; et sur F_2 : $F(45, 882) = 11,06$; $p < 0,001$.

Afin de mesurer l'effet des variations inter-individuelles et contextuelles sur la dispersion des voyelles **prototypiques** identifiées en AM, nous avons calculé la taille des ellipses de dispersion pour chaque voyelle en fonction du contexte de présentation. Les résultats présentés sur la Figure 120 ne montrent pas d'effets systématiques de variations de taille des ellipses en fonction des **prototypes** identifiés. Nous observons par contre que les trois voyelles cardinales /i: a: u:/ obtiennent les ellipses les plus petites et plus particulièrement pour /i:/, les ellipses les plus grandes sont obtenues pour les trois voyelles brèves /ɪ ɘ ʊ/. Il semblerait qu'en **dynamique** les effets de la densité des systèmes vocaliques n'affectent pas la taille des ellipses des trois voyelles extrêmes /i: a: u:/, conformément aux propositions de la **théorie quantique**.

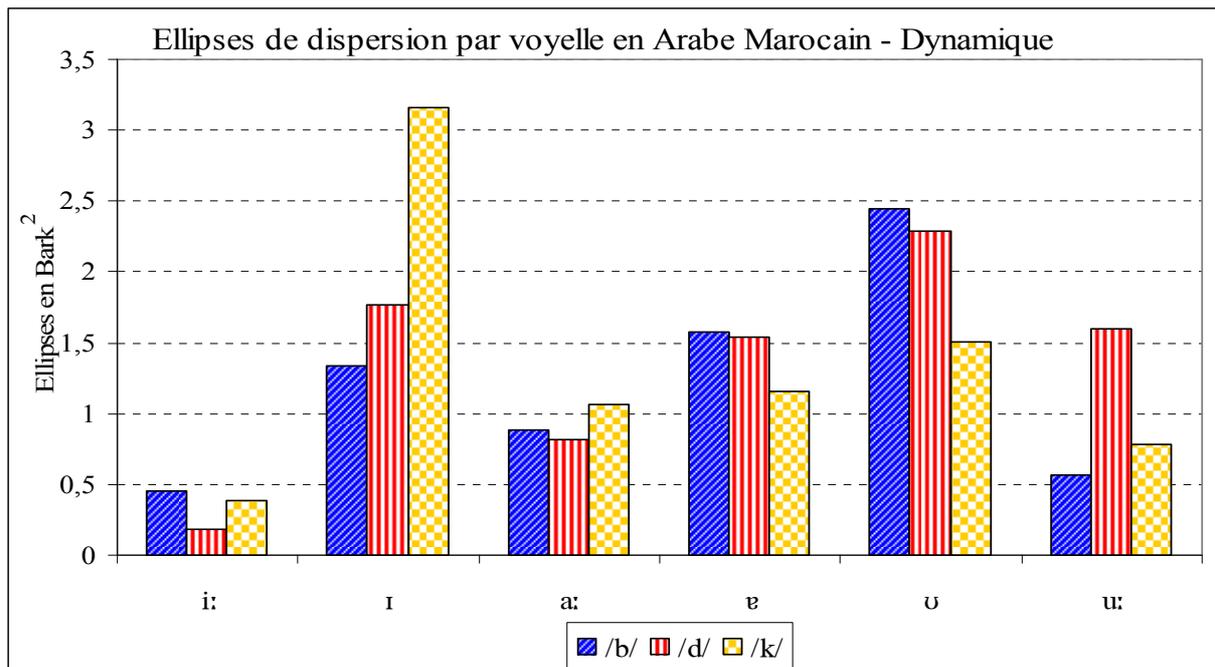


Figure 120 : Ellipses de dispersion par voyelle **prototypique** perçue en AM en **dynamique**, en fonction des contextes de présentation.

Ces résultats montrent que les auditeurs marocains étaient sensibles aux variations observées au niveau de la qualité des six timbres vocaliques **prototypiques** et au niveau des influences consonantiques sur la position de ces voyelles. Nous retrouvons une homogénéité entre les résultats obtenus en production et en perception. Nous évaluerons dans la partie qui suit les résultats obtenus pour la comparaison entre les **prototypes** vocaliques identifiés en **statique** et en **dynamique**.

5.3.2.3. Comparaison des résultats en statique vs en dynamique

Les résultats de la comparaison entre les **prototypes** vocaliques obtenus en **statique** vs **dynamique** dans le choix des **prototypes** vocaliques en AM, montrent des différences significatives uniquement sur F_2 : les voyelles /I ə a: ʊ u:/ sont significativement plus « antériorisées » en **dynamique** qu'en **statique** ($F(5, 1764)=40,26$; $p<0,001$), (voir Figure 121). Cette différence est plus marquée pour les voyelles /I ə a:/ que pour les voyelles postérieures ; aucune différence n'est obtenue pour la voyelle /i:/. La dispersion des voyelles à l'intérieur de l'espace vocalique semble différente en fonction des traits utilisés, mais sans

modification de la taille des aires de dispersion : 17,28 Bark² en **statique** et 17,62 en **dynamique** (différence non significative : $F(1, 18)=0,81$; $p=0,42$).

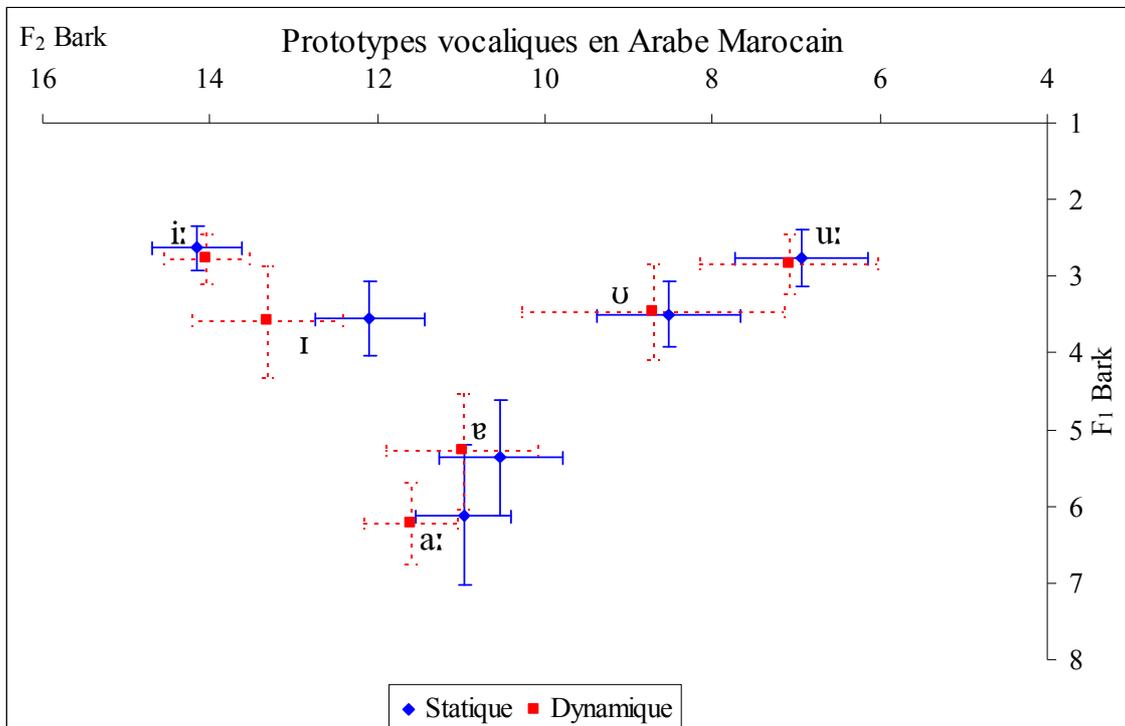


Figure 121 : Espaces vocaux des **prototypes** identifiés en AM en **statique** vs en **dynamique**.

Les résultats obtenus par Johnson (2000) ont montré que l'utilisation de la **dynamique** (i.e., consonne initiale ajoutée à la voyelle) n'a pas affecté la position des voyelles **prototypiques** identifiées par les auditeurs anglais, ni la taille des aires de dispersion. Nos résultats sur l'AM sont partiellement en accord avec ceux de Johnson (2000) pour ce qui est de la taille des espaces vocaux non significativement différente. En revanche, nos résultats montrent qu'à l'inverse des propositions de Johnson (2000), les **prototypes** vocaliques choisis par les auditeurs marocains sont significativement différents en **statique** vs en **dynamique**. Autrement dit, les auditeurs de l'AM sont sensibles aux effets coarticulatoires de la consonne initiale sur la voyelle.

Les résultats de l'interaction entre les traits utilisés (**statique** vs **dynamique**) et les trois contextes de présentation montrent des différences significatives sur les deux axes. Sur F_1 , un effet non systématique est obtenu, car en contexte /b/ par exemple, l'espace vocalique est significativement plus « ouvert » en **dynamique**, tandis que c'est en **statique** que l'espace

est plus « ouvert » en contexte dental ($F(2, 1764)= 7,32$; $p<0,001$) ; des différences très marquées sont obtenues en fonction des timbres vocaliques (nous y reviendrons un peu plus loin). Sur F_2 , les trois espaces vocaliques obtenus en **dynamique** et dans les trois contextes de présentation sont significativement plus « antériorisées » qu'en **statique** ($F(2, 1764)=14,15$; $p<0,001$), (voir Figure 122). L'interaction entre les traits utilisés (**statique** vs **dynamique**), le contexte de présentation et les voyelles donne des différences significatives sur les deux axes. Ainsi, sur F_1 , les voyelles /i: ɐ a: ʊ/ en contexte bilabial sont significativement plus « ouvertes » et /ɪ/ est plus « fermée » en **dynamique** ; en contexte dental, les voyelles /i: ʊ u:/ sont plus « ouvertes » et /ɪ ɐ a:/ sont plus « fermées » en **dynamique** ; et en contexte vélaire, /ɪ u:/ sont significativement plus « ouvertes » et /ɐ a: ʊ/ sont plus « fermées » en **dynamique**, ($F(10, 1764)= 5,04$; $p<0,001$). Sur F_2 , les voyelles /ɪ a: ʊ/ en contexte bilabial sont significativement plus « antérieures » et /i: u:/ sont plus « postérieures » en **dynamique** ; en contexte dental, les voyelles /ɪ ɐ a: ʊ u:/ sont significativement plus « antérieures » et seule /i:/ est plus « postérieure » en **dynamique** ; et enfin en contexte vélaire, /ɪ ɐ a: u:/ sont significativement plus « antérieures » et /ʊ/ est plus « postérieure » en **dynamique**, ($F(10, 1764)= 10,47$; $p<0,001$).

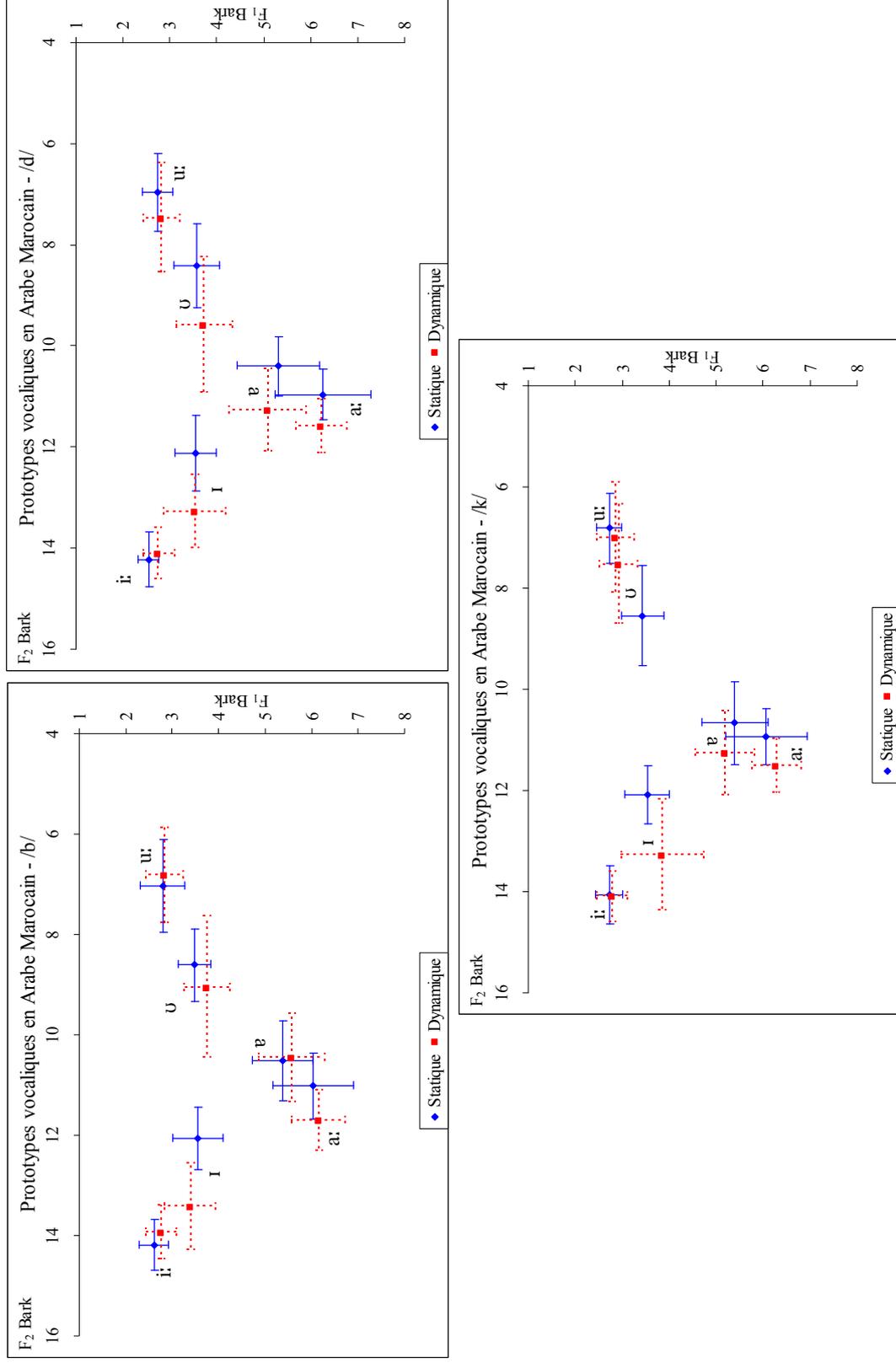


Figure 122 : Espaces vocales des prototypes en AM en statique vs dynamique, en fonction des contextes de présentation.

Nous avons voulu vérifier si les différences observées sur les trois graphiques de la Figure 122, en ce qui concerne la taille des aires de dispersion d'une part et de la position des voyelles, d'autre part, sont statistiquement significatives. Pour cela, nous avons calculé la taille des aires de dispersion des voyelles par contexte de présentation en fonction des traits utilisés et effectué des analyses statistiques complémentaires pour chaque voyelle (T-test de Student : test d'égalité des espérances : observations appariées). Sur la Figure 123, nous présentons la taille des aires de dispersion par contexte de présentation et par trait utilisé : les aires de dispersion des voyelles **prototypes** identifiées en AM ne présentent pas de taille statistiquement différente, quels que soient les contextes de présentation et les modalités.

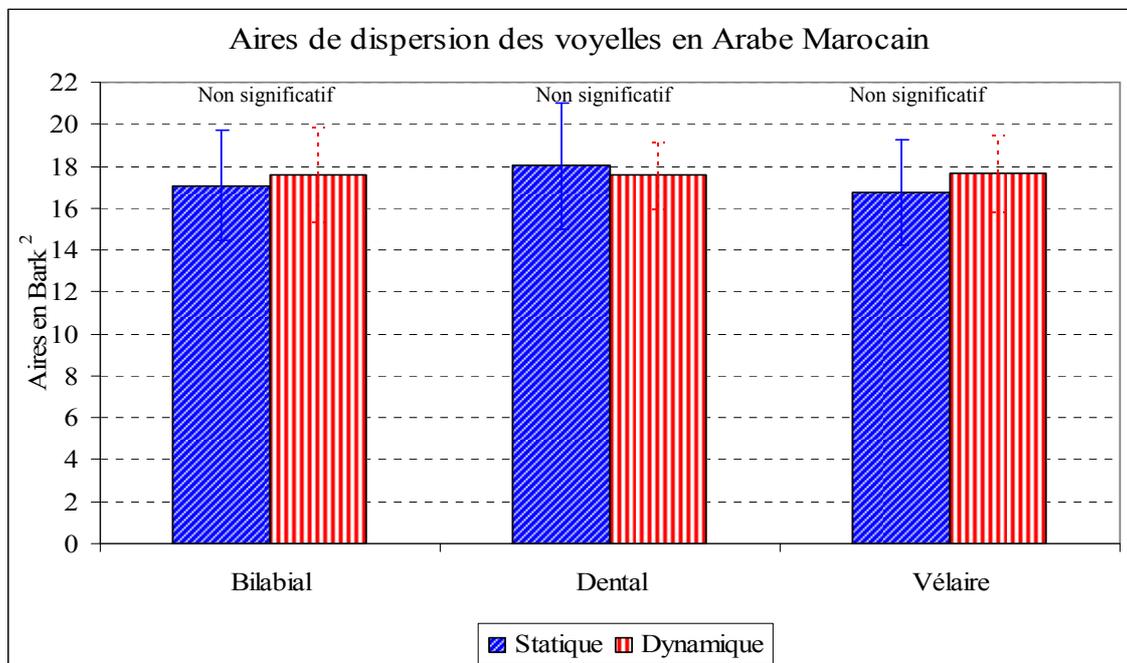


Figure 123 : Aires de dispersion des voyelles **prototypiques** identifiées en AM en **statique** vs **dynamique**, en fonction des contextes de présentation (ns = non significatif).

Pour ce qui est des différences de positions des voyelles **prototypiques** présentées sur la Figure 122, les résultats statistiques, présentés dans la Tableau 39, montrent que les différences les plus marquées sont obtenues pour la voyelle /*v*/ dans les environnements /*d k*/ sur les deux axes F_1 et F_2 ; pour la voyelle /*a*/ sur F_2 dans les trois contextes de présentation ; et pour la voyelle /*u*/ dans l'environnement vélaire sur les deux axes F_1 et F_2 .

Tableau 39 : Résultats des T-test de Student entre les voyelles **prototypiques** identifiées par les auditeurs marocains, en **statique** vs **dynamique**. [+F] = [+fermé], [+O] = [+ouvert], [+A] = [+antérieur], [+P] = [+postérieur], NS = Non Significative. Le tableau se lit de la façon suivante : /i/ > /b/ > F₁ = [+F], p<0,05 : la voyelle /i/ identifiée en contexte /b/ en **dynamique** est significativement [+fermée] (p<0,05) par rapport à celle en **statique**.

	F ₁			F ₂		
	/b/	/d/	/k/	/b/	/d/	/k/
/i/	[+F], p<0,05	NS	[+O], p<0,05	[+A], p<0,001	[+A], p<0,001	[+A], p<0,001
/i:/	[+O], p<0,01	[+O], p<0,001	NS	[+P], p<0,01	NS	NS
/e/	[+O], p<0,05	[+F], p<0,01	[+F], p<0,01	NS	[+A], p<0,001	[+A], p<0,001
/a:/	NS	NS	[+O], p<0,01	[+A], p<0,001	[+A], p<0,001	[+A], p<0,001
/o/	[+O], p<0,001	[+O], p<0,05	[+F], p<0,001	[+A], p<0,05	[+A], p<0,001	[+P], p<0,001
/u:/	NS	NS	[+O], p<0,05	[+P], p<0,05	[+A], p<0,001	NS

Les différences de taille des aires de dispersion et de la position des voyelles **prototypiques** identifiées par les auditeurs marocains sont dépendantes des stratégies individuelles : il existe en effet une variabilité inter-individuelle importante qui est relevée sur les deux axes et quantifiée par les écart-types et par la taille des ellipses de chaque voyelle. Ainsi, nous présentons sur la Figure 124 les ellipses de dispersion par voyelle en fonction des traits utilisés. Ces différences ne sont pas significatives, $F(1, 10)=0,005$; $p=0,95$. Aucune tendance n'est observée à l'issue de cette analyse, sauf pour les deux voyelles fermées /i/ u:/, dont les ellipses de dispersion en **statique** et en **dynamique** sont plus petites comparées aux autres voyelles ; de plus, les ellipses des trois voyelles brèves sont les plus grandes par rapport aux autres voyelles (sauf la voyelle /a:/ en **statique**). Nous obtenons les mêmes résultats en comparant la taille des ellipses de dispersion par catégorie vocalique en fonction du contexte de présentation et des traits utilisés (i.e., aucune différence significative, voir Annexe 33).

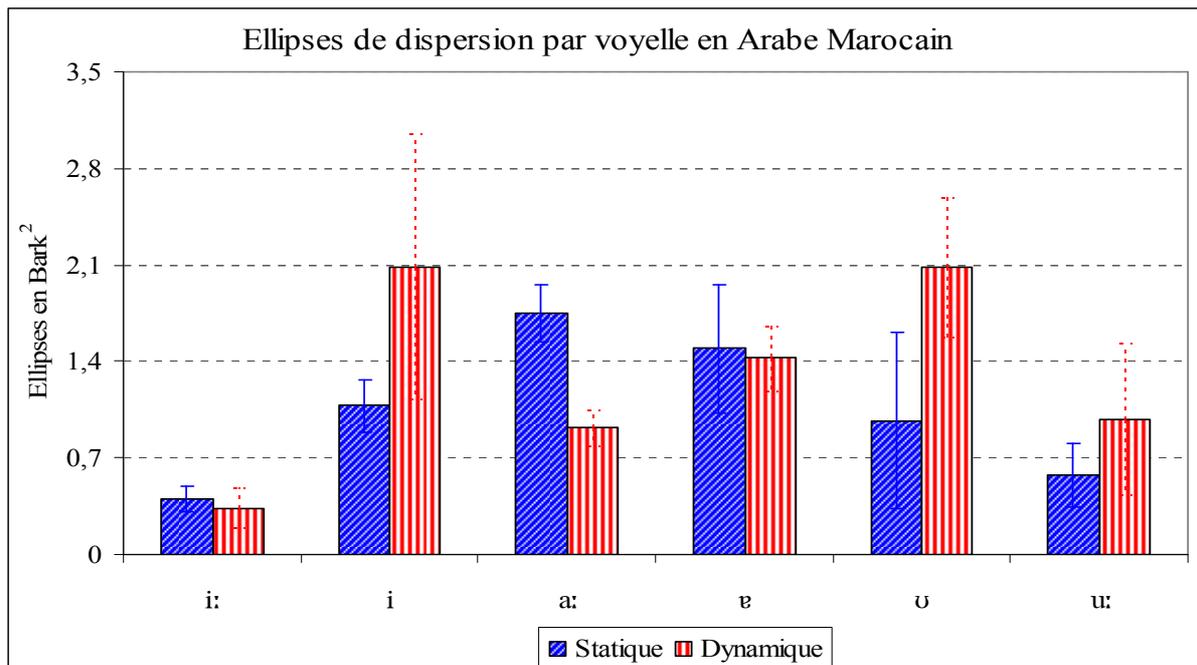


Figure 124 : Ellipses de dispersion par voyelle **prototypique** perçue en AM en **statique** vs **dynamique**.

Tous les auditeurs marocains, participant à ces deux expériences (i.e., recherche des **prototypes** vocaliques en **statique** et en **dynamique**), nous ont expliqué que la tâche de perception en **statique** était moyennement difficile par rapport à celle en **dynamique**. Les explications avancées mettent en cause la qualité des sons de synthèse (i.e., sons correspondant à des « *bips* » ou des « *bruits* ») ou encore l'absence de variation dans la structure temporelle et fréquentielle de la voyelle (i.e., des voyelles « *trop soutenues* » ou « *trop longues* »). Tous les auditeurs marocains nous ont expliqué que les aspects **dynamiques** (i.e., « *les consonnes avant* ») ont « *énormément* » aidés à reconnaître les voyelles et à les identifier.

Les résultats des **prototypes** vocaliques identifiés par les auditeurs marocains ont montré, d'une part, que les traits **dynamiques** apportent plus de précision et de « *naturel* » aux stimuli synthétiques lors de l'identification des voyelles en AM, et d'autre part que les effets du contexte de présentation sont visibles surtout en **dynamique**. Des différences qualitatives ont été observées entre les voyelles **prototypiques** en AM. Nous avons ensuite évalué la validité de nos résultats par une analyse discriminante.

5.3.2.4. Résultats de l'analyse discriminante⁷⁵

Nous avons appliqué une analyse discriminante avec validation croisée, ayant comme paramètres les valeurs de F_1 et de F_2 des **prototypes** vocaliques identifiés par les auditeurs marocains. La classification a été proposée en fonction des voyelles identifiées en **statique** vs **dynamique** et du contexte de présentation. Nous avons montré que les timbres des voyelles **prototypiques** identifiés par les auditeurs marocains étaient significativement différents en **statique** et en **dynamique** et quel que soit le contexte de présentation. Les taux de classification de l'analyse discriminante confirment ces résultats (voir Tableau 40).

Tableau 40 : Taux de classification correcte des voyelles **prototypiques** identifiés en AM en **statique** vs **dynamique** en fonction de chaque contexte de présentation. (Taux significatifs, Box's M ; $p < 0,001$).

	/b/	/d/	/k/
Statique	85,3%	86,3%	85,7%
Dynamique	79,0%	79,0%	74,7%

L'observation des taux de classification révèle que la **dynamique** semble ne pas améliorée la séparation entre les voyelles. Nous obtenons un taux moyen de classification des voyelles en **dynamique** inférieur de 8% à celui en **statique** (BinomialQ ; $p < 0,001$)⁷⁶. Nous ne nous attendions pas à obtenir un taux moyen plus bas en **dynamique**, car les auditeurs marocains ont expliqué que la tâche d'identification des voyelles était plus facile lorsque les

⁷⁵ Le taux de classification de l'analyse discriminante est considéré comme significatif si l'analyse statistique Box's M, proposée dans l'Analyse, donne des résultats significatifs. Le Box's M est une analyse statistique qui propose de valider (ou non) l'hypothèse nulle selon laquelle la variance/covariance pour la totalité de l'échantillon est identique à travers les données. Autrement dit, s'il existe des différences entre les sous-groupes testés, la moyenne et la variance vont être significativement différentes. La valeur de Box's M doit être la plus élevée possible : plus la valeur de Box's M est élevée, plus le seuil de significativité est bas (i.e., la probabilité est inférieure à 0,001). Si le seuil de significativité est supérieur à 0,05, l'analyse discriminante n'est pas valide (i.e., on ne peut pas effectuer une classification des données). Dans ce cas, nous mentionnerons ces taux et tenterons de les expliquer.

⁷⁶ Le BinomialQ (p, k, n) est une analyse statistique, proposée dans Praat, qui permet de calculer la probabilité pour que, dans un nombre précis d'expérience n , un événement avec une probabilité p intervienne avec un nombre de fois k . Dans notre étude, nous calculons ainsi la probabilité pour que le taux de classification obtenu en **statique** p , soit significativement différent du taux de classification en **dynamique**. La méthode de calcul utilisée est la suivante : BinomialQ (p, k, n) = BinomialQ (taux de classification en **statique**, taux de classification en **dynamique** * la somme du nombre total de voyelles dans les deux conditions, la somme du nombre total de voyelles dans les deux conditions).

traits **dynamiques** étaient utilisés. Il semblerait toutefois que ces traits **dynamiques** n'aient pas amélioré la perception des auditeurs marocains, à l'inverse de leurs suppositions. Nous avons vu dans les parties précédentes que l'identification des **prototypes** en AM était variable en fonction des auditeurs. La plupart des confusions dans la classification des voyelles étaient moins importantes lorsque les traits **statiques** ont été utilisés (voir Figure 125) plutôt que les traits **dynamiques** (voir Figure 126).

L'examen de ces confusion montre que les trois voyelles longues /i: a: u:/ identifiées par les auditeurs marocains en **statique** ont été classifiées avec des taux allant de 78% à 98%, quel que soit le contexte de présentation. La plupart des confusions obtenues concernait les deux voyelles /a: ɐ/ qui sont très proches (voir également Figure 115), et qui ont fait baisser les taux de classification finaux. Les deux voyelles /ɪ ʊ/ ont été souvent bien classifiées, mais les confusions obtenues sont causées par leurs positions proches des voyelles /i: u:/.

Les trois voyelles longues /i: a: u:/ identifiées par les auditeurs marocains en **dynamique** ont été classifiées avec des taux allant de 68% à 96%. Les confusions observées concernaient soit les deux voyelles /u: ʊ/ qui sont très proches, soit les couples de voyelle /a: ɐ/ et /i: ɪ/, soit les trois voyelles brèves /ɪ ɐ ʊ/ (qui sont très proches selon les réalisations individuelles, voir la Figure 118).

Il ressort de cette classification le constat suivant : les trois voyelles **prototypiques** /ɪ ɐ ʊ/ identifiées par les auditeurs marocains sont toutes les trois bien classifiées et très peu confondues entre elles (les confusions étaient toujours avec les voyelles correspondantes longues). Les auditeurs marocains ont choisi trois « timbres » vocaliques brèves comme « représentant » de la catégorie /ə/.

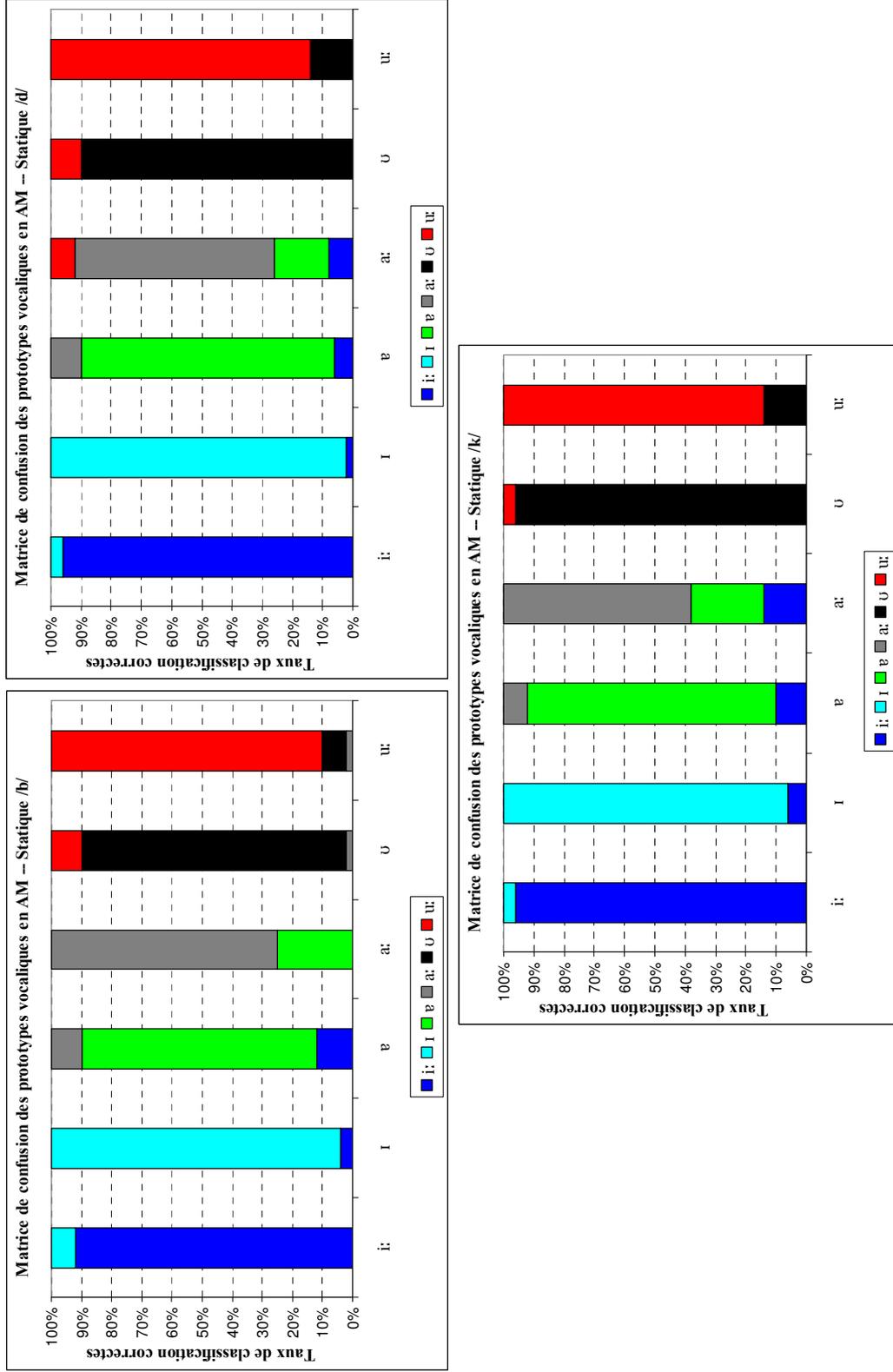


Figure 125 : Matrice de confusion des prototypes en AM en statique, en fonction des contextes de présentation.

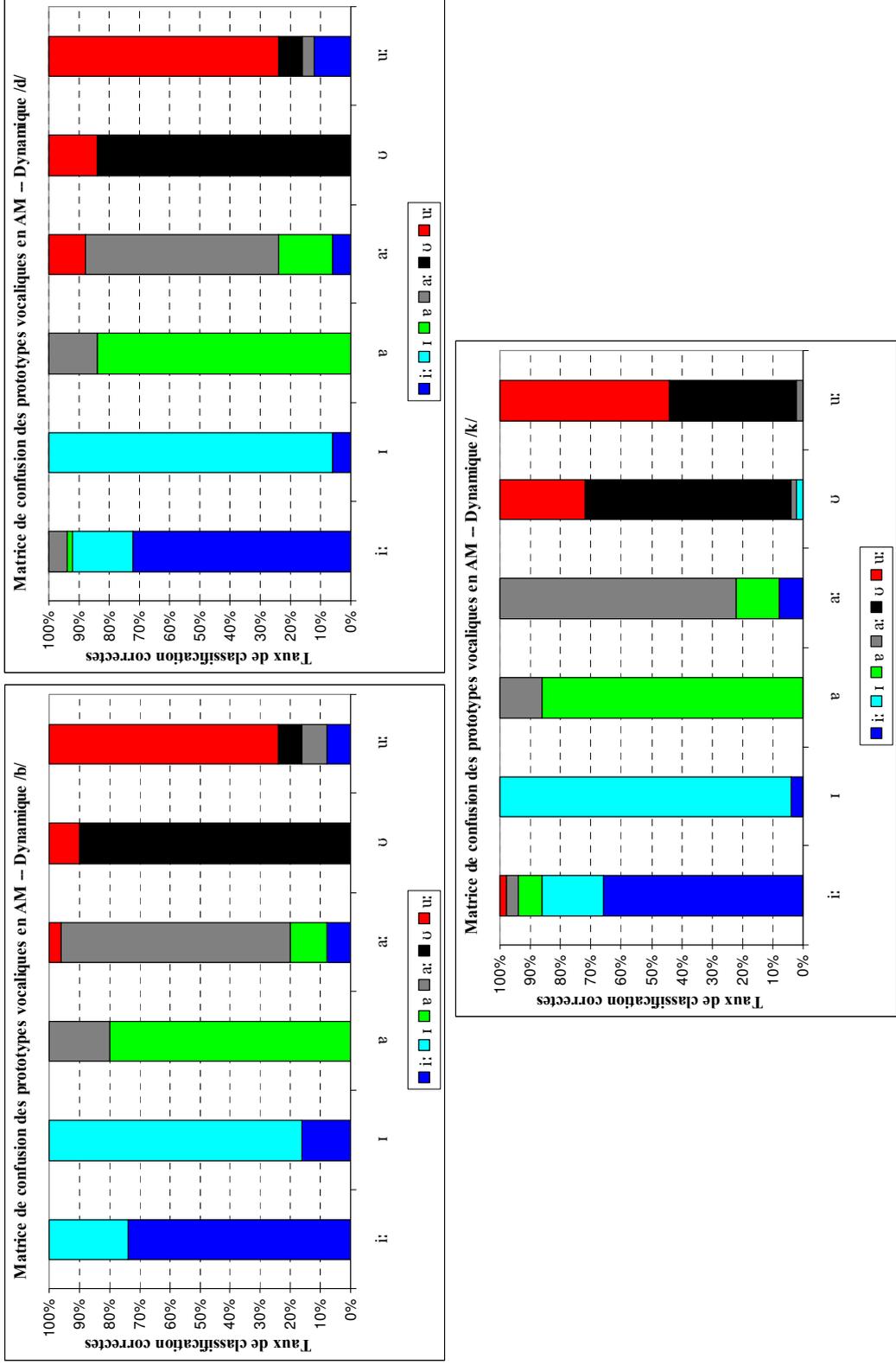


Figure 126 : Matrice de confusion de classification des prototypes en AM en dynamique, en fonction des contextes de présentation

Nous nous posons des questions sur le comportement des auditeurs marocains : pourquoi observons-nous trois timbres vocaliques brefs /ɪ ɐ ʊ/ identifiés comme **prototypes** vocaliques perceptifs ? Pourquoi les **prototypes** vocaliques perceptifs choisis par les auditeurs marocains ne correspondent pas aux résultats de la littérature ? Est-ce que les auditeurs marocains sont conscients des différences qualitatives entre ces trois voyelles brèves et le font ressortir sur le plan perceptif ? Ou la structure syllabique a-t-elle influencé la perception des auditeurs marocains (la plupart des items utilisés en AM sont monosyllabiques, par conséquent toutes les voyelles se trouvent accentuées, et donc sont réalisées comme des voyelles « non réduites ») ? Nous pensons en effet que l'utilisation d'items monosyllabiques en AM peut affecter la réalisation des voyelles réduites et plus particulièrement celles en **dynamique**, car la perception des effets coarticulatoires des consonnes sur les voyelles peut expliquer ces différences. Néanmoins, en **statique**, aucune information autre que la structure acoustique de la voyelle n'est fournie dans les stimuli ; comment les auditeurs marocains ont-ils fait pour choisir trois **prototypes** vocaliques non réduits différents comme « représentant » de la catégorie vocalique /ə/ ? Les taux de classification et de confusion obtenus par l'analyse discriminante confirment cette tendance. Nous supposons que la « représentation mentale » différente pour ces voyelles a affecté la perception des auditeurs marocains. Ainsi, leurs choix sont fonction du statut des voyelles brèves dans leur système. Ces résultats nous poussent à proposer un système vocalique perceptif en AM composé de six timbres vocaliques. Nous examinerons plus en détails le statut perceptif de ces voyelles lors de la seconde expérience de perception : la catégorisation des espaces vocaliques (voir § 5.4.2).

Intéressons-nous maintenant aux résultats de perception des **prototypes** en AJ.

5.3.3. *En Arabe Jordanien*

Nous présentons ici les résultats obtenus en perception des **prototypes** vocaliques par les auditeurs jordaniens. Nous commençons par présenter les résultats en **statique** ; ensuite en **dynamique** et terminons par les résultats de la comparaison entre les traits utilisés et les résultats de l'analyse discriminante.

5.3.3.1. Résultats en statique (V Target)

Les **prototypes** vocaliques identifiés en AJ en **statique** sont présentés sur la Figure 127. Comme en AM, nous n'observons aucune influence globale du contexte de présentation sur les espaces vocaliques obtenus en AJ, ni sur F_1 : $F(2, 1174) = 0,66$; $p=0,52$; ni sur F_2 : $F(2, 1617) = 2,56$; $p=0,1$. Ces résultats vont dans les sens des propositions de la littérature en ce qui concerne l'homogénéité de l'espace vocalique perceptif quel que soit le contexte de présentation, car comme nous l'avons expliqué, cet espace perceptif est stable : la perception de la parole est un phénomène *many-to-one* où plusieurs stimuli acoustiques sont identifiés comme un seul et unique stimulus auditif (voir également p. 315).

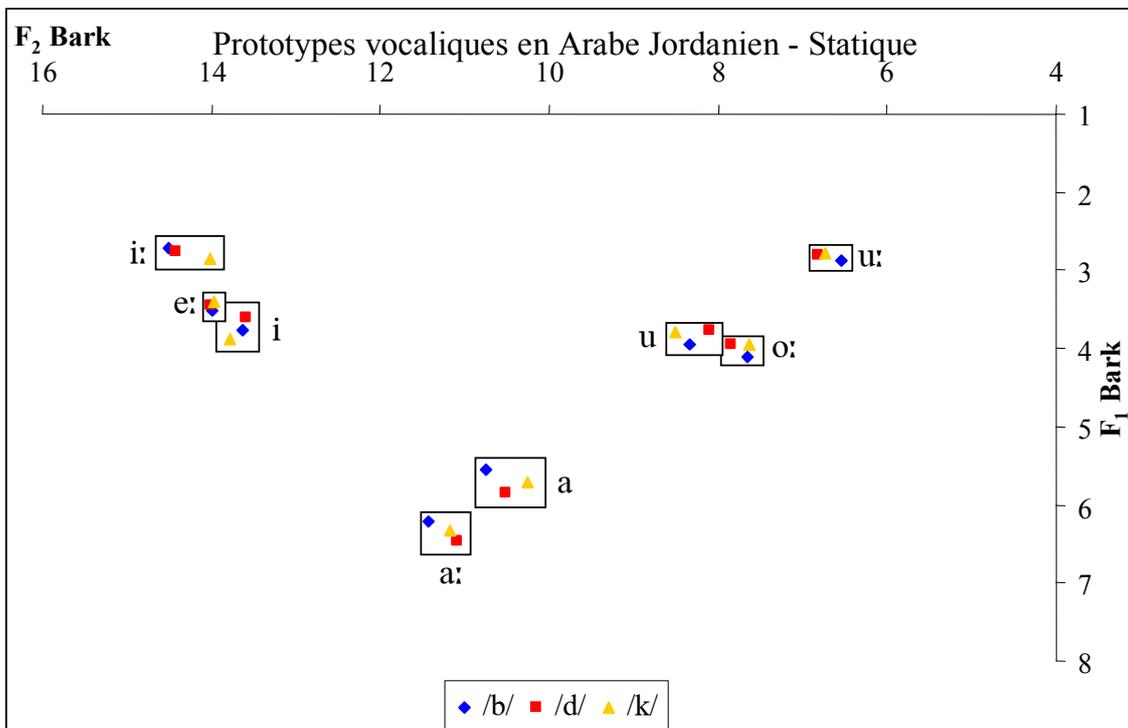


Figure 127 : Espace vocalique des **prototypes** vocaliques identifiés en AJ en **statique**, en fonction des contextes de présentation.

En ce qui concerne l'interaction entre les contextes de présentation et les voyelles, les résultats ne montrent aucune différence ni sur F_1 ($F(14, 1174) = 1,38$; $p=0,15$), ni sur F_2 : $F(14, 1174) = 1,14$; $p=0,32$. Ceci indique qu'en **statique**, les espaces vocaliques **prototypiques** identifiés par les auditeurs jordaniens ne sont pas significativement différentes, quels que soient les voyelle et les contexte des présentation, allant ainsi dans le

sens des propositions de Meunier (2005). Nous n'avons trouvé aucune différence significative quant à la taille des aires de dispersion en fonction du contexte de présentation (voir Tableau 41).

Tableau 41 : Taille des aires de dispersion des voyelles **prototypiques** identifiées en AJ en **statique**, en fonction des contextes de présentation (en Bark²).

	Moyenne	Écart-Type
Bilabial	16,06	6,50
Dental	16,56	6,40
Vélaire	16,09	6,41

Nous observons également des différences significatives entre les auditeurs jordaniens (voir Figure 128) sur F_1 : $F(9, 1174) = 915,59$; $p < 0,001$, et sur F_2 : $F(9, 1174) = 1733,97$; $p < 0,001$. Les trois voyelles longues /i: a: u:/ sont choisies dans les extrémités de l'espace vocalique perceptif ; les deux voyelles périphériques /e: o:/ retrouvent leurs « positions » sur le triangle vocalique et les trois voyelles brèves /i a u/ se trouvent dans une position « centrale » dans l'espace acoustico-perceptif. Nous observons le rapprochement des voyelles brèves /i/ et /u/ des voyelles /e:/ et /o:/ respectivement, i.e., des voyelles [e] et [o] ; en ce qui concerne les deux voyelles ouvertes /a: a/, nous obtenons des différences significatives quant à leurs positions sur l'espace acoustico-perceptif, conformément aux résultats obtenus en production (sur F_1 : $F(9, 1174) = 23,83$; $p < 0,001$ et sur F_2 : $F(9, 1174) = 35,19$; $p < 0,001$). L'interaction entre les auditeurs et les voyelles est significative : sur F_1 : $F(63, 1174) = 8,30$; $p < 0,001$; et sur F_2 : $F(63, 1174) = 9,87$; $p < 0,001$.

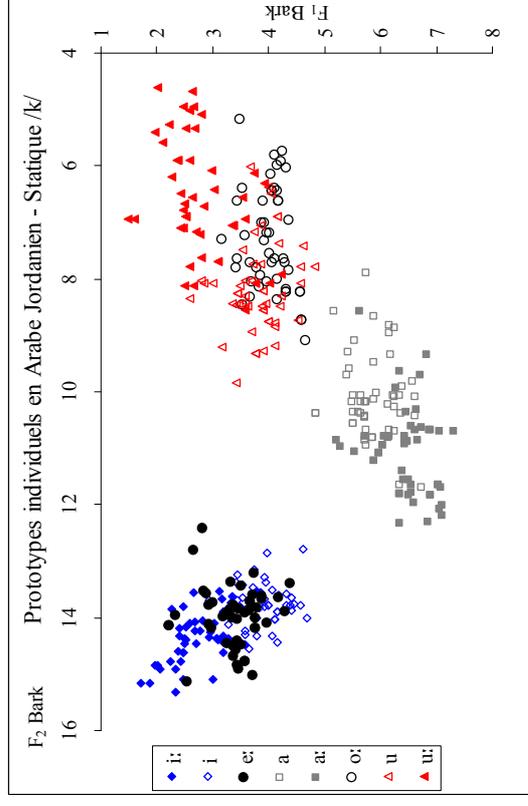
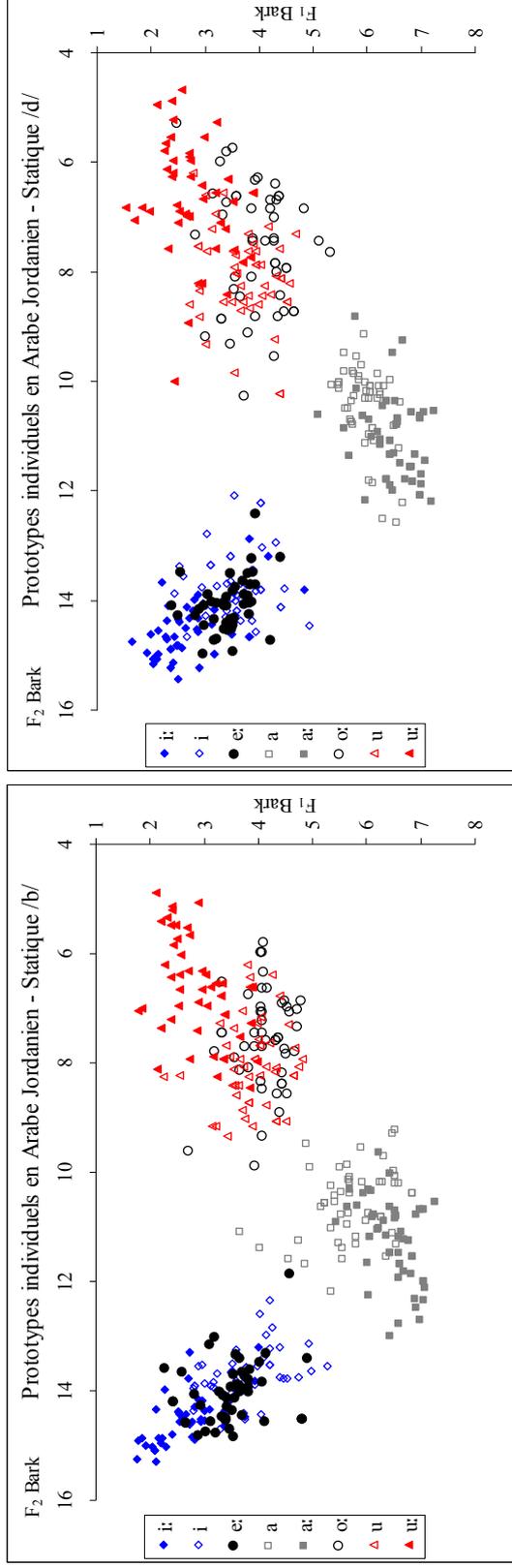


Figure 128 : Dispersion des voyelles **prototypiques** individuelles en AJ en **statique**, en fonction des contextes de présentation.

Afin de quantifier ces différences, avec la représentation graphique, nous avons calculé la taille des ellipses de dispersion par voyelle, en fonction du contexte de présentation (voir Figure 129).

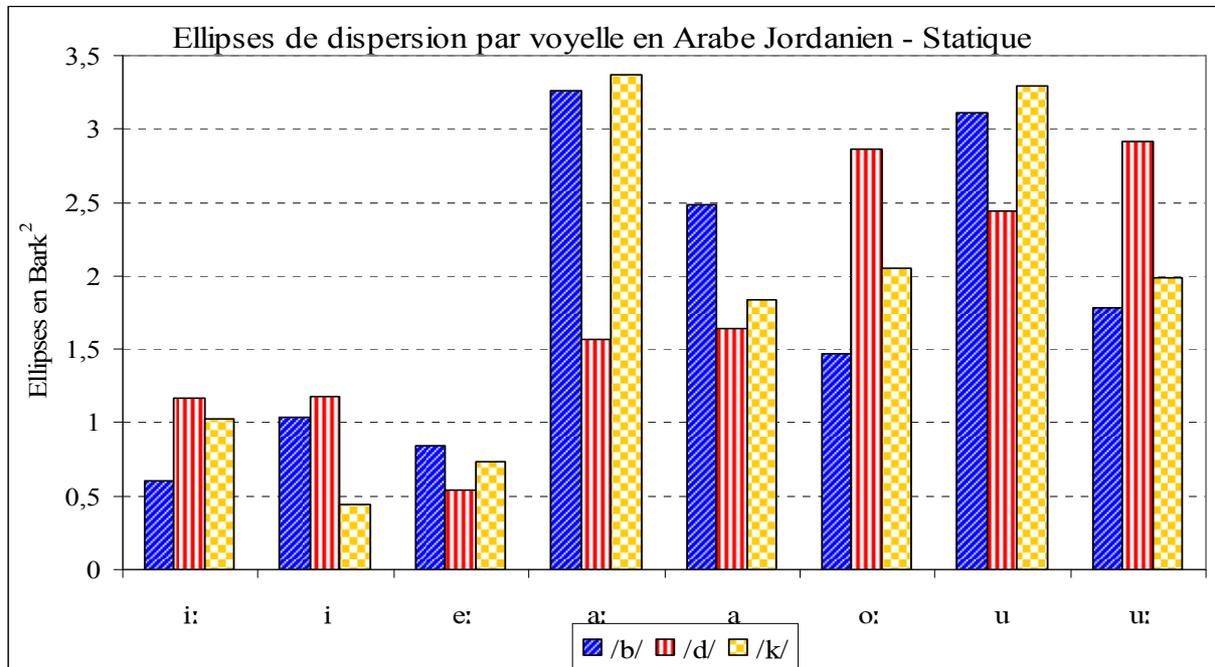


Figure 129 : Ellipses de dispersion par voyelle **prototypique** identifiée en AJ en **statique**, en fonction des contextes de présentation.

Les résultats graphiques de la dispersion des voyelles au niveau individuel (Figure 128) et de la taille des ellipses par voyelle (Figure 129) ne montrent pas d'effets systématiques. Nous observons des ellipses plus grandes pour les voyelles « centrales » et « postérieures » par rapport aux voyelles « antérieures », montrant une variabilité inter-individuelle et contextuelle importante. Nous pourrions corrélés ces résultats aux impressions des auditeurs jordaniens après le test. En effet, tous les auditeurs jordaniens avaient trouvé cette tâche très difficile. Nous supposons donc que la taille des ellipses de chaque voyelle très grande est sensiblement corrélée à la confusion dans leurs choix.

5.3.3.2. Résultats en dynamique (CV Target)

Les résultats graphiques des voyelles **prototypiques** identifiées en AJ en **dynamique** sont présentés sur la Figure 130. Le contexte de présentation semble influencer le choix des auditeurs jordaniens et ce sur F_1 uniquement, $F(2, 1176) = 10,24$; $p < 0,001$; aucun effet sur

F_2 , $F(2, 960) = 1,25$; $p=0,29$ n'est détecté. Cette différence observée sur F_1 va à l'encontre de l'hypothèse de stabilité de l'espace vocalique perceptif (Meunier, 2005). Les auditeurs jordaniens choisissent des voyelles significativement différentes sur l'axe fermé~ouvert (F_1).

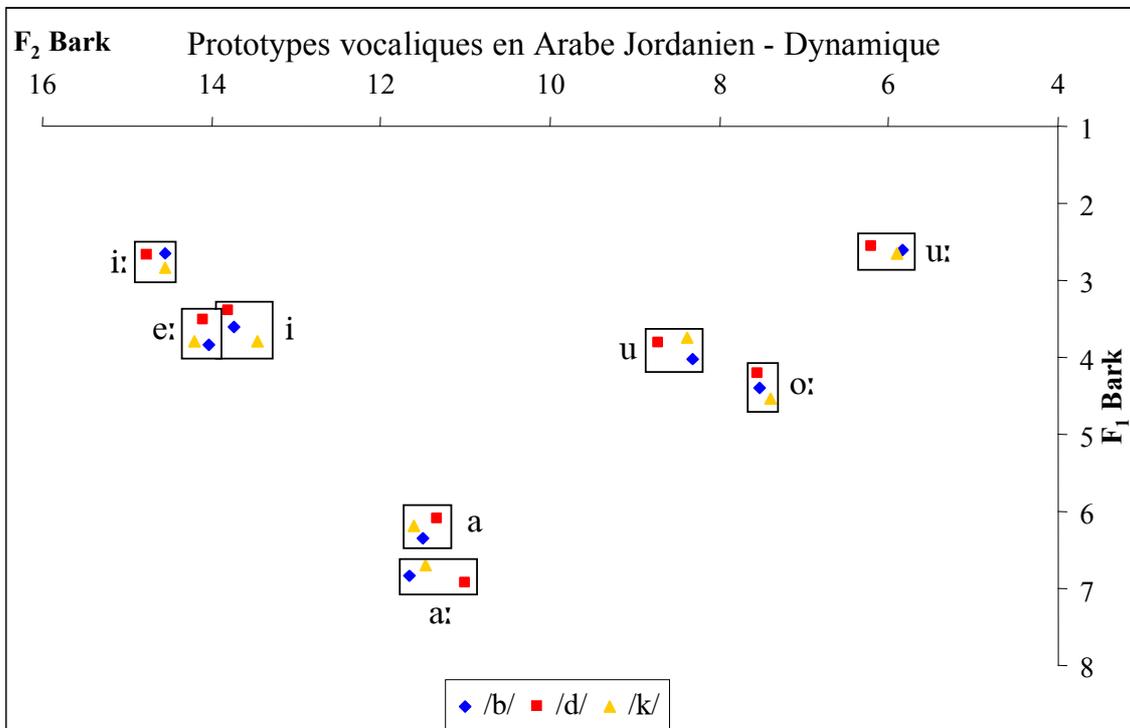


Figure 130 : Espace vocalique des **prototypes** vocaliques identifiés en AJ en **dynamique**, en fonction des contextes de présentation.

Comme en **statique**, nous observons une interaction significative entre les voyelles et les contextes de présentation sur les deux axes : sur F_1 : $F(14, 1176)=3,51$; $p<0,001$; sur F_2 : $F(14, 1176)=3,69$; $p<0,001$. Toutefois, cet effet n'est pas systématique, car sur F_1 , les voyelles sont plus « fermées » tantôt dans le contexte dental (e.g., [i e: a o: u:]), tantôt dans le contexte vélaire (e.g., [a: u]) ; et sur F_2 , les voyelles sont plus « antériorisées » tantôt dans le contexte vélaire (e.g., [e: a]), tantôt dans le contexte dental (e.g., [i: i u u:]), tantôt dans le contexte bilabial (e.g., [a:]).

Afin de quantifier ces différences, nous avons calculé la taille des ellipses de dispersion des voyelles en fonction des contextes de présentation, dont les résultats sont présentés dans le Tableau 42. Les résultats montrent que la taille des aires de dispersion en **dynamique** semble être identique quel que soit le contexte de présentation (avec néanmoins

une différence moyenne non significative de ≈ 1 Bark² entre les espaces vocaliques obtenus en contexte bilabial et en contexte vélaire).

Tableau 42 : Taille des aires de dispersion des voyelles **prototypiques** identifiées en AJ en **dynamique**, en fonction des contextes de présentation (en Bark²)

	Moyenne	Écart-Type
Bilabial	21,06	3,93
Dental	20,73	4,54
Vélaire	20,20	4,71

L'étude de la dispersion des voyelles (voir Figure 130) montre que les auditeurs jordaniens distinguent les voyelles longues aux brèves en termes de timbres vocaliques différents. Nous obtenons ainsi des différences significatives entre les timbres des voyelles **prototypiques** identifiées par les auditeurs jordaniens : sur F_1 , $F(7, 1176) = 1654,09$; $p < 0,001$; sur F_2 , $F(7, 1176) = 3205,61$; $p < 0,001$. Comme en **statique**, les voyelles longues sont positionnées dans les extrémités de l'espace vocalique ; les brèves quant à elles ont des comportements différents. Ainsi, les deux voyelles brèves /i u/ sont positionnées aux côtés des voyelles longues [e: o:], respectivement, et non pas aux côtés de leurs correspondantes longues [i: u:]. En ce qui concerne la voyelle /a/, elle a été identifiée par les auditeurs jordaniens comme significativement différente de sa correspondante longue [a:].

Il existe des différences significatives liées aux choix des auditeurs jordaniens (voir Figure 131) sur F_1 , $F(9, 1176) = 23,33$; $p < 0,001$ et sur F_2 , $F(9, 1176) = 9,79$; $p < 0,001$, et en fonction de l'interaction entre les auditeurs et les voyelles sur F_1 , $F(63, 1176) = 5,84$; $p < 0,001$ et sur F_2 , $F(63, 1176) = 12,73$; $p < 0,001$.

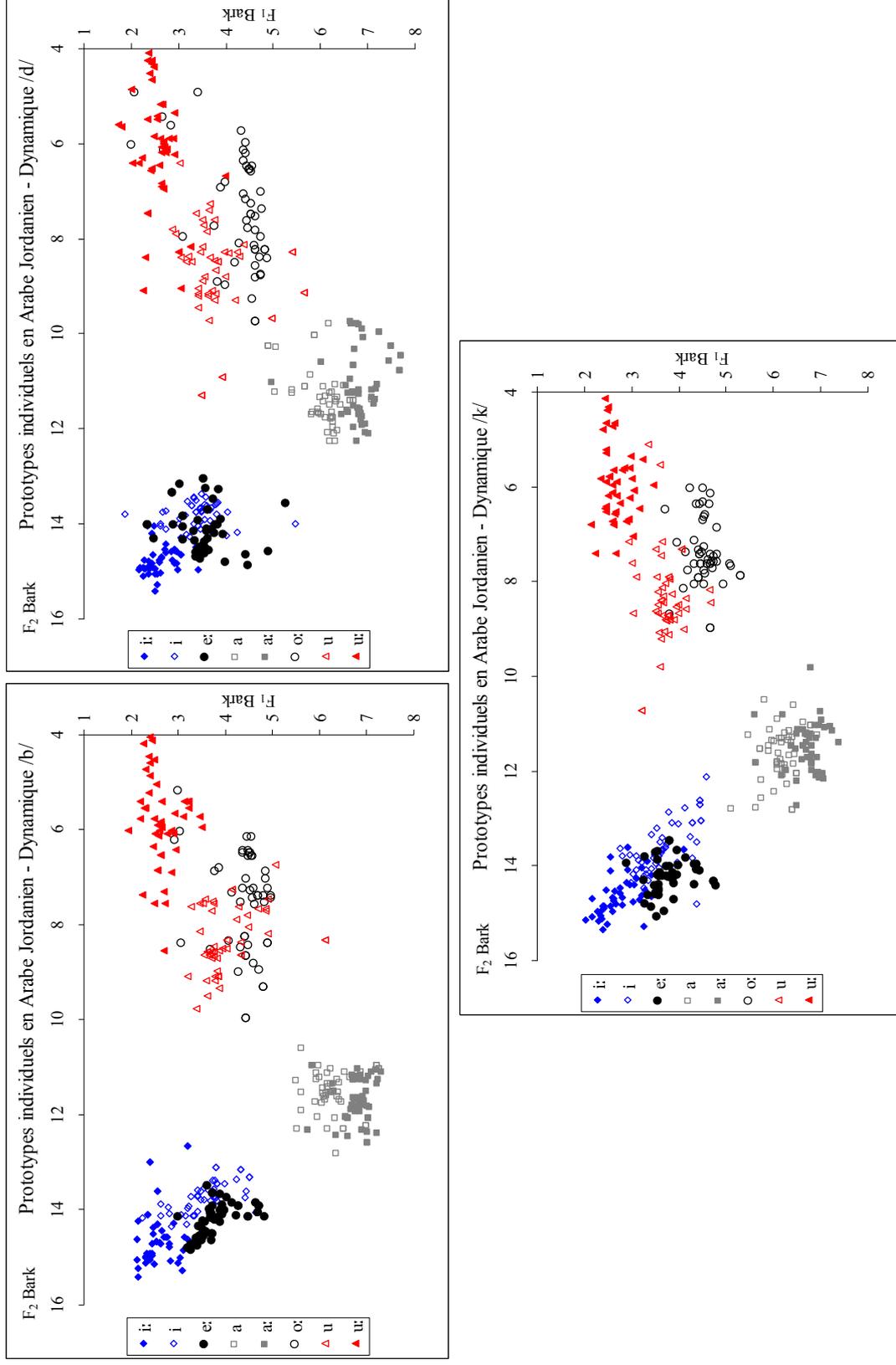


Figure 131 : Dispersion des voyelles **prototypiques** individuelles en AJ en **dynamique**, en fonction des contextes de présentation.

Nous avons quantifié cette variabilité en fonction des contextes de présentation à partir du calcul de la taille des ellipses de dispersion par voyelle. Ainsi, les résultats graphiques présentés sur la Figure 132 montrent des variations importantes en fonction de la taille des ellipses des voyelles identifiées par les auditeurs jordaniens en fonction des différents contextes de présentation. Dans certains cas, c'est en contexte dental que les ellipses de dispersion sont les plus grandes (e.g., [e: a: o: u]), tandis que c'est en contexte bilabial que l'ellipse de la voyelle [i] est la plus grande. Les voyelles antérieures obtiennent des ellipses de dispersion plus petites par rapport aux postérieures (différences non significatives).

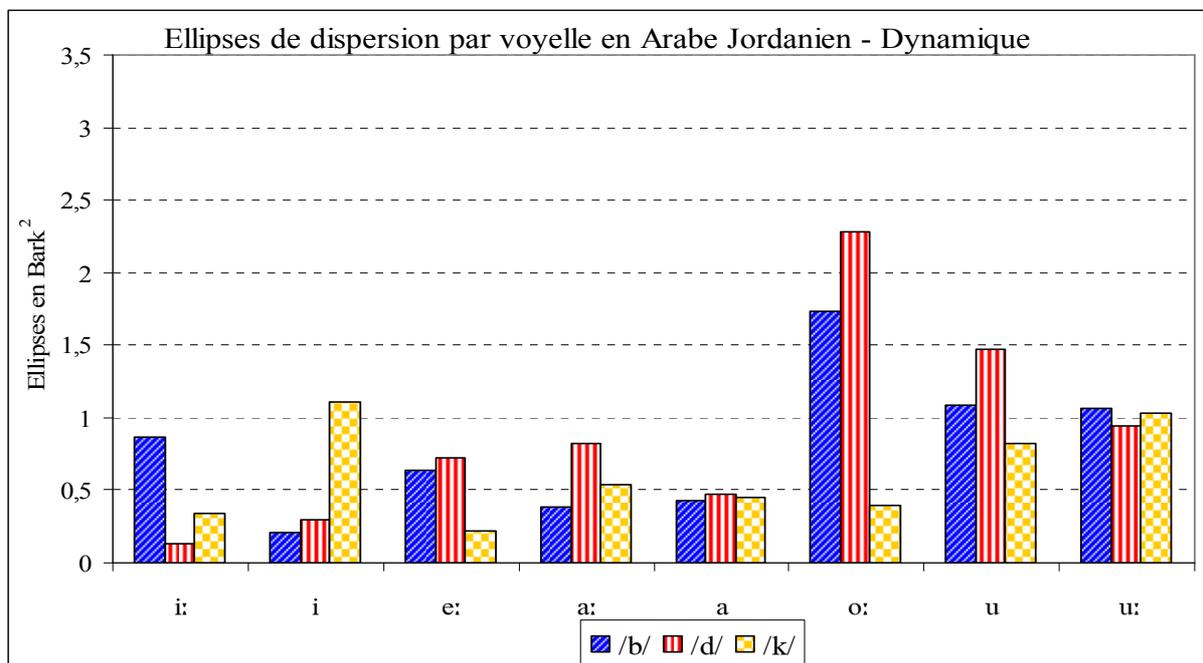


Figure 132 : Ellipses de dispersion par voyelle **prototypique** perçue en AJ en **dynamique**, en fonction des contextes de présentation.

Les résultats obtenus en perception des voyelles **prototypiques** en AJ montrent une homogénéité entre les résultats obtenus en production et en perception, d'un côté, et entre les résultats en perception en **statique** et en **dynamique**, de l'autre. Nous retrouvons ainsi les voyelles brèves /i u/ positionnées aux côtés des voyelles longues /e: o:/ respectivement et non pas aux côtés de leurs correspondantes longues comme décrit par la littérature (voir § 1.4.1).

5.3.3.3. Comparaison des résultats en statique vs en dynamique

Nous comparons dans cette partie les résultats des choix des **prototypes** vocaliques en AJ en **statique** et en **dynamique**. Les résultats graphiques montrent des différences significatives entre les deux espaces vocaliques obtenus ; l'espace vocalique perceptif obtenu en **dynamique** est significativement plus grand que celui observé en **statique** et ce sur les deux axes F_1 ($F(1, 2350) = 35,33$; $p < 0,001$) et F_2 ($F(1, 2350) = 31,58$; $p < 0,001$), (voir Figure 133). La taille de l'aire de dispersion des voyelles identifiées en **dynamique** est significativement plus grande que celle en **statique** (16,24 Bark² en **statique** contre 20,66 en **dynamique** : $F(1, 14) = 220,48$; $p < 0,001$).

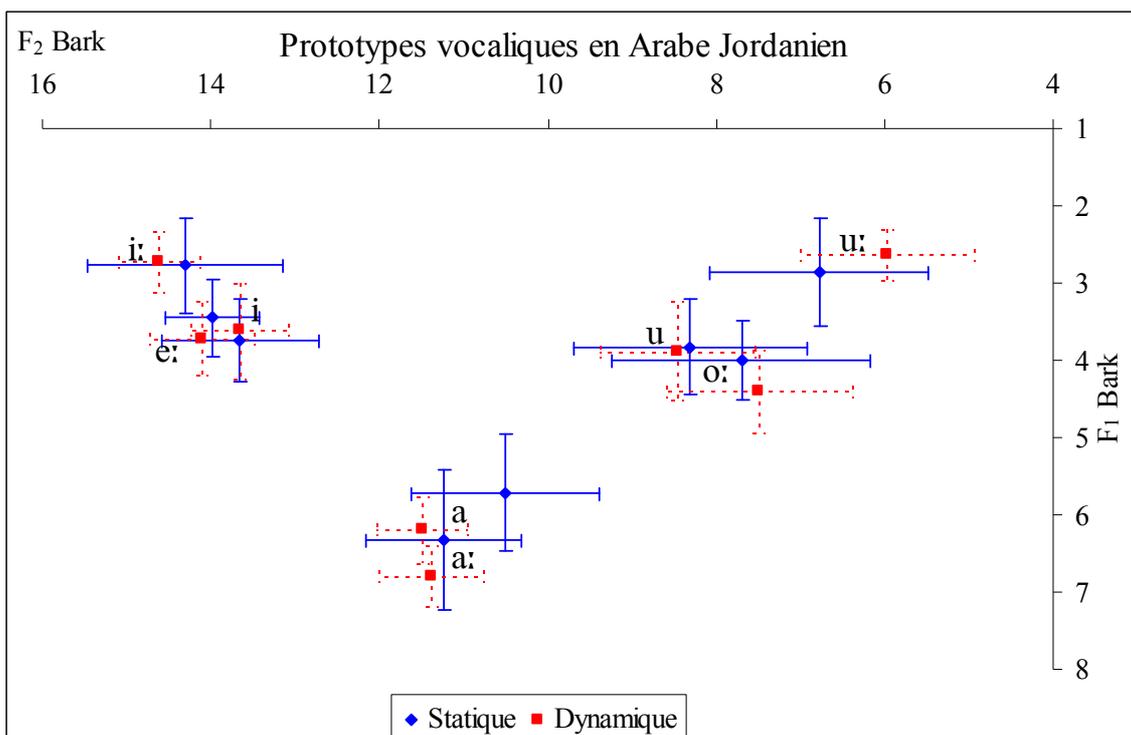


Figure 133 : Espaces vocaliques des **prototypes** AJ en **statique** vs **dynamique**.

Ce résultat va à l'encontre des propositions de Johnson (2000), selon qui aucune différence significative n'est obtenue pour ce qui est de la taille des espaces vocaliques des **prototypes** identifiés en **dynamique**. Nous ne nous attendions pas à obtenir cette différence entre les espaces vocaliques car selon les propositions de Meunier (2005), le phénomène *many-to-one* qui explique l'organisation des espaces vocaliques perceptifs, aide les auditeurs

à « normaliser » les différences existant entre les stimuli acoustiques variables afin de retrouver un stimulus auditif commun.

En ce qui concerne les voyelles problématiques /i a u/, nos résultats obtenus en **statique** et en **dynamique** montrent que la voyelle /i/ est différente de sa correspondante longue /i:/ et est identifiée comme [e] ; la voyelle /u/ est différente de /u:/ et est identifiée comme [o] ; et enfin, /a/ est plus « fermée » que sa correspondante longue /a:/ et est identifiée comme [ɐ]. L'interaction entre les traits **statiques** et **dynamiques** et les voyelles est statistiquement significative. Ainsi, en **dynamique**, sur F_1 , les voyelles /i: i u:/ sont plus « fermées » et /e: a a: o:/ sont plus « ouvertes » qu'en **statique** ($F(7, 2350) = 16,53$; $p < 0,001$). Sur F_2 , les voyelles /i: e: a a: u/ sont significativement plus « antériorisées » et les voyelles /o: u:/ sont significativement plus « postérieures » en **dynamique** par rapport aux voyelles identifiées en **statique** ($F(7, 2350) = 35,50$; $p < 0,001$).

Les différences observées en fonction des traits utilisés (Figure 133) sont homogènes quels que soient le trait utilisé et le contexte de présentation (voir la Figure 134). Les résultats graphiques montrent les mêmes tendances : un espace vocalique des **prototypes** identifiés par les auditeurs jordaniens plus grand en **dynamique** par rapport à l'espace identifié en **statique** quel que soit le contexte de présentation. Les résultats statistiques de cette interaction ne montrent aucune différence significative, autrement dit, les tendances observées sont identiques quel que soit le contexte de présentation : sur F_1 , $F(2, 2350) = 2,32$; $p = 0,1$; F_2 , $F(2, 2350) = 0,71$; $p = 0,5$.

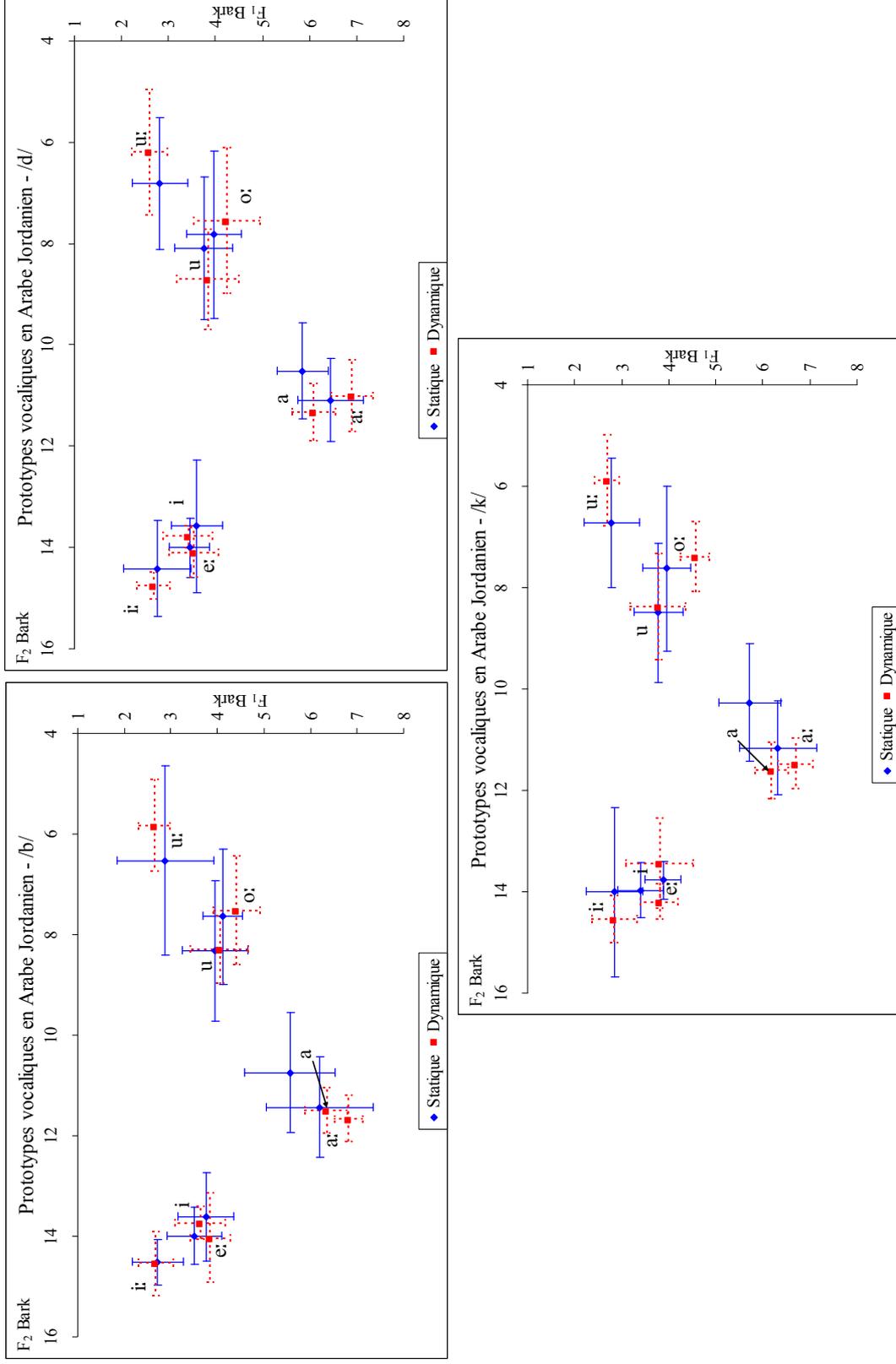


Figure 134 : Espaces vocaliques des prototypes en AJ en statique vs dynamique, en fonction des contextes de présentation.

Pour vérifier ces résultats, nous avons calculé la taille de dispersion des voyelles identifiées en fonction des traits utilisés et des contextes de présentation. Les résultats graphiques présentés sur la Figure 135 montrent que la taille des aires en **dynamique** est, en moyenne, 27% plus grande par rapport à la taille des aires en **statique** : effet significatif en contexte bilabial (31% plus grand en **dynamique** pour le contexte bilabial ; $p < 0,05$) et une tendance est obtenue pour les contextes dental et vélaire (25% plus grand en **dynamique** pour les contextes dental et vélaire ; $p = 0,06$). Ces différences sont faiblement significatives à cause des résultats individuels sont variables. Nous obtenons en effet des différences de réponses entre les auditeurs jordaniens assez marquées : l'auditeur AJM03 présente des aires de dispersion très petites en **statique** (3,69 Bark² en /b/ ; 6,02 Bark² en /d/ ; et 6,19 Bark² en /b/) et grandes en **dynamique** (19,49 Bark² en /b/ ; 20,86 Bark² en /d/ ; et 18,05 Bark² en /b/) ; tandis que l'auditeur AJM10 présente des aires de dispersion des voyelles *presque* identiques entre les deux modalités : en **statiques** (26,52 Bark² en /b/ ; 25,93 Bark² en /d/ ; et 29,29 Bark² en /b/) et en **dynamique** (28,71 Bark² en /b/ ; 28,44 Bark² en /d/ ; et 29,44 Bark² en /b/).

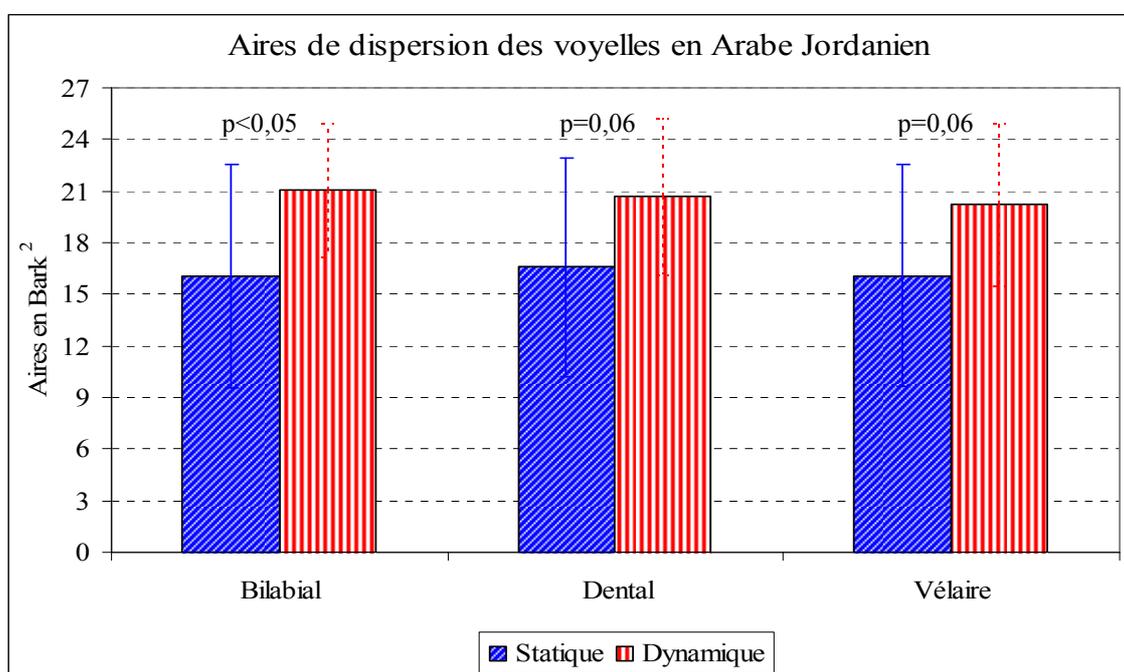


Figure 135 : Aires de dispersion des voyelles **prototypiques** identifiées en AJ en **statique** vs **dynamique**, en fonction des contextes de présentation.

Les différences inter-individuelles ne se limitent pas à la différence globale de la taille de dispersion des espaces vocaliques perceptifs ; nous observons des différences importantes

liées à la taille des ellipses de chaque voyelle en **statique** et en **dynamique**. Les résultats graphiques présentés sur la Figure 136 montrent que l'utilisation des traits **dynamiques** apporte plus de facilité aux auditeurs jordaniens en leur donnant plus de précision pour le choix des **prototypes** vocaliques de leur système. Ceci se traduit par des ellipses significativement plus petites ($F(1, 14)= 10,25 ; p<0,007$), (voir Annexe 34 pour les résultats en fonction des contextes de présentation). L'utilisation de la **dynamique** a permis aux auditeurs jordaniens de choisir des **prototypes** vocaliques significativement éloignés des mêmes **prototypes** en **statique**, i.e., nous obtenons un effet d'*Hyper-Articulation* (Lindblom, 1990) de l'espace vocalique en **dynamique** et un effet d'*Hypo-Articulation* en **statique**. L'*Hyper-Articulation* de l'espace se traduit par des positions très extrêmes des **prototypes** dans l'espace acoustico-perceptif, mais également une précision des choix. Lindblom (1990) explique que lorsque les voyelles sont en mode *Hyper-Articulé*, la dispersion intra-catégorie vocalique est très faible et c'est notre cas, car les auditeurs jordaniens choisissent leurs prototypes dans les extrémités de l'espace vocalique et la taille des ellipses de chaque voyelle est très faible.

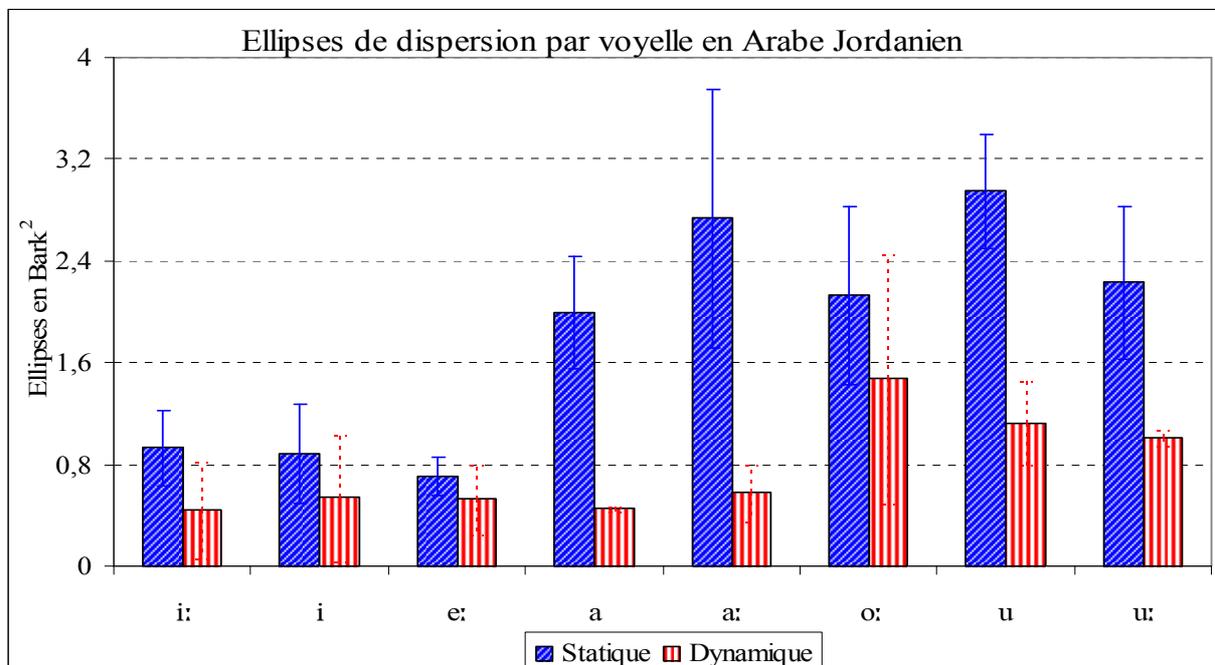


Figure 136 : Aires de dispersion par voyelle **prototypique** perçue en AJ en **statique** vs **dynamique**.

Les résultats statistiques montrent l'existence d'un effet global des différences observées au niveau de la position des voyelles **prototypiques** identifiées en AJ en **statique**

vs **dynamique**, en fonction des voyelles mêmes (sur F_1 ($F(1) = 69,9$; $p < 0,001$) ; sur F_2 ($F(1) = 9,07$; $p < 0,003$)) et des interactions entre les voyelles et les traits utilisés (sur F_1 : $F(7) = 30,97$; $p < 0,001$; sur F_2 : $F(7) = 34,51$; $p < 0,001$), entre les voyelles et le contexte de présentation (sur F_1 : $F(14) = 4,70$; $p < 0,001$; sur F_2 : $F(14) = 3,06$; $p < 0,001$) et enfin en fonction de l'interaction entre les voyelles, les traits utilisés et le contexte de présentation (sur F_1 : $F(14) = 2,55$; $p < 0,002$; sur F_2 : $F(14) = 3,10$; $p < 0,001$). C'est au niveau de cette dernière interaction que nous allons détailler les analyses.

En effet, nous voulons savoir si les positions des voyelles **prototypiques** identifiées par les auditeurs jordaniens sont significativement différentes (ou non). Ainsi, un T-test de Student : test d'égalité des espérances : observations appariées, a été appliqué aux données. Les résultats sont présentés dans le Tableau 43. D'une façon générale, les résultats montrent que les voyelles **prototypiques** occupent les mêmes positions dans les espaces vocaliques perceptifs de l'AJ et plus spécifiquement pour les deux voyelles brèves /i u/ qui ne sont pas identifiées différemment entre les deux modalités (i.e., en **statique** ou en **dynamique**).

Tableau 43 : Résultats des T-test de Student entre les voyelles **prototypiques** identifiées par les auditeurs jordaniens, en **statique** vs **dynamique**. [+F] = [+fermé], [+O] = [+ouvert], [+A] = [+antérieur], [+P] = [+postérieur], NS = Non Significative. Le tableau se lit de la façon suivante : /i/ > /d/ > F_1 = [+F], $p < 0,05$: la voyelle /i/ identifiée en contexte /d/ en **dynamique** est [+fermée] ; $p < 0,05$, par rapport à celle en **statique**.

	F ₁			F ₂		
	/b/	/d/	/k/	/b/	/d/	/k/
/i/	NS	[+F], $p < 0,05$	NS	NS	NS	[+A], $p < 0,01$
/i:/	NS	NS	NS	NS	[+A], $p < 0,01$	[+A], $p < 0,05$
/e:/	[+O], $p < 0,001$	NS	[+O], $p < 0,001$	NS	NS	[+A], $p < 0,005$
/a/	[+O], $p < 0,001$	[+O], $p < 0,05$	[+O], $p < 0,001$	[+A], $p < 0,001$	[+A], $p < 0,001$	[+A], $p < 0,001$
/a:/	[+O], $p < 0,001$	[+O], $p < 0,001$	[+O], $p < 0,005$	[+A], $p < 0,05$	NS	[+A], $p < 0,05$
/o:/	[+O], $p < 0,001$	[+O], $p < 0,005$	[+O], $p < 0,001$	NS	NS	NS
/u/	NS	NS	NS	NS	[+A], $p < 0,005$	NS
/u:/	[+F], $p < 0,01$	[+F], $p < 0,005$	NS	[+P], $p < 0,001$	[+P], $p < 0,005$	[+P], $p < 0,001$

Comme pour les auditeurs marocains, la tâche d'identification des **prototypes** vocaliques en AJ en **statique** a posé beaucoup de problèmes aux auditeurs jordaniens. Ces auditeurs ont expliqué que les sons qu'ils entendaient ne correspondaient pas du tout à des sons de parole, mais plutôt à des « bips » ou des « bruits » d'ordinateur. Ils expliquaient que

les sons avec les « *consonnes à l'initiale* » étaient aussi difficiles à rapprocher des sons de parole, mais qu'ils facilitaient l'identification des voyelles. Les résultats que nous avons obtenus dans ce travail et les réponses des auditeurs jordaniens aux questions d'après tests vont dans le sens d'une amélioration de l'identification des **prototypes** vocaliques en AJ. En **dynamique**, les auditeurs jordaniens étaient sensibles aux variations contextuelles apportées par l'environnement consonantique ajouté à la voyelle. Tous les auditeurs jordaniens ont choisi des **prototypes** vocaliques pour les trois voyelles brèves /i a u/ différents de ceux des voyelles longues correspondantes. Plus spécifiquement, les auditeurs jordaniens ont choisi des **prototypes** significativement différents entre les couples de voyelles /i u/ et /i: u:/, respectivement : des différences significatives très importantes sont obtenues entre /i/ et /i:/ ($p=1^{E-9}$) ; et entre /u/ et /u:/ ($p=1^{E-9}$) ; les différences existant entre les couples de voyelles /i/ et /e:/ et entre /u/ et /o:/ sont beaucoup plus faibles à cause du rapprochement entre elles. Ceci montre que les différences qualitatives obtenues en production des voyelles sont confirmées par les auditeurs jordaniens. Il serait judicieux de vérifier la validité de nos résultats et de valider l'apport des traits **dynamiques** dans la séparation entre les voyelles en AJ, ce avec l'aide de l'analyse discriminante dont nous présentons les résultats dans la partie qui suit.

5.3.3.4. Résultats de l'analyse discriminante

Les paramètres utilisés pour cette analyse sont les valeurs de F_1 et de F_2 des **prototypes** vocaliques choisis par les auditeurs jordaniens. Les taux de classification des voyelles **prototypiques** identifiées en AJ en fonction des traits utilisés et des contextes de présentation, sont présentés dans le Tableau 44. L'analyse discriminante a permis de séparer les voyelles de l'AJ avec des taux significatifs proches des 70%. L'utilisation de la **dynamique** a facilité la tâche des auditeurs jordaniens lors de l'identification des **prototypes** de leur système. Le taux moyen de classification des voyelles obtenu en **dynamique** est supérieur de 15% à celui en **statique** (BinomialQ ; $p<0,001$), confirmant ainsi nos résultats.

Tableau 44 : Taux de classification correcte des voyelles **prototypiques** identifiées en AJ en **statique** vs **dynamique** en fonction des contextes de présentation. (Taux significatifs, Box's M ; $p < 0,001$).

	/b/	/d/	/k/
Statique	61,1%	60,5%	67,3%
Dynamique	76,3%	76,0%	82,0%

Bien que nous obtenions une amélioration du taux de classification de l'analyse discriminante de près de 15%, ces taux sont bas. Ceci est dû sans doute à la non-utilisation de la durée comme variable, mais également aux confusions dans la classification. Étant donnée la proximité des deux voyelles brèves /i u/ des deux voyelles longues /e: o:/ mise en évidence dans nos résultats, nous nous attendons à observer le maximum de confusions entre ces couples de voyelles.

Ainsi, nous observons sur les graphiques de la Figure 137, les taux de confusion pour la classification des voyelles **prototypiques** identifiées par les auditeurs jordaniens en **statique**, par contexte de présentation. Les trois voyelles longues /i: a: u:/ ont été classifiées correctement avec un taux ne dépassant pas les 84%. La grande majorité des confusions concernait les deux voyelles brèves /i u/ qui étaient souvent confondues avec les deux voyelles longues /e: o:/, respectivement. Dans un seul cas, en contexte /d/, la voyelle /u/ a été mal classifiée, car dans 46% des cas, cette voyelle a été classifiée comme une voyelle /o:/ et 32% seulement comme /u/ (les 22% restant comme /u:/). Ceci indique la proximité des deux voyelles (i.e., elles ont les mêmes caractéristiques et l'analyse discriminante les a confondues).

Nous observons l'augmentation du taux de classification des voyelles **prototypiques** identifiées par les auditeurs jordaniens en **dynamique** et ainsi la baisse du taux de confusion de classification. Les résultats graphiques, présentés sur la Figure 138, montrent que toutes les voyelles longues obtiennent des taux de classification très élevés. Surtout, toutes les occurrences de la voyelle /u:/ en contexte vélaire sont classifiées à 100% comme /u:/. La plupart des erreurs de classification sont obtenues pour les voyelles brèves et plus spécifiquement /i u/ qui sont confondues, comme en **statique**, avec les deux voyelles longues /e: o:/, respectivement.

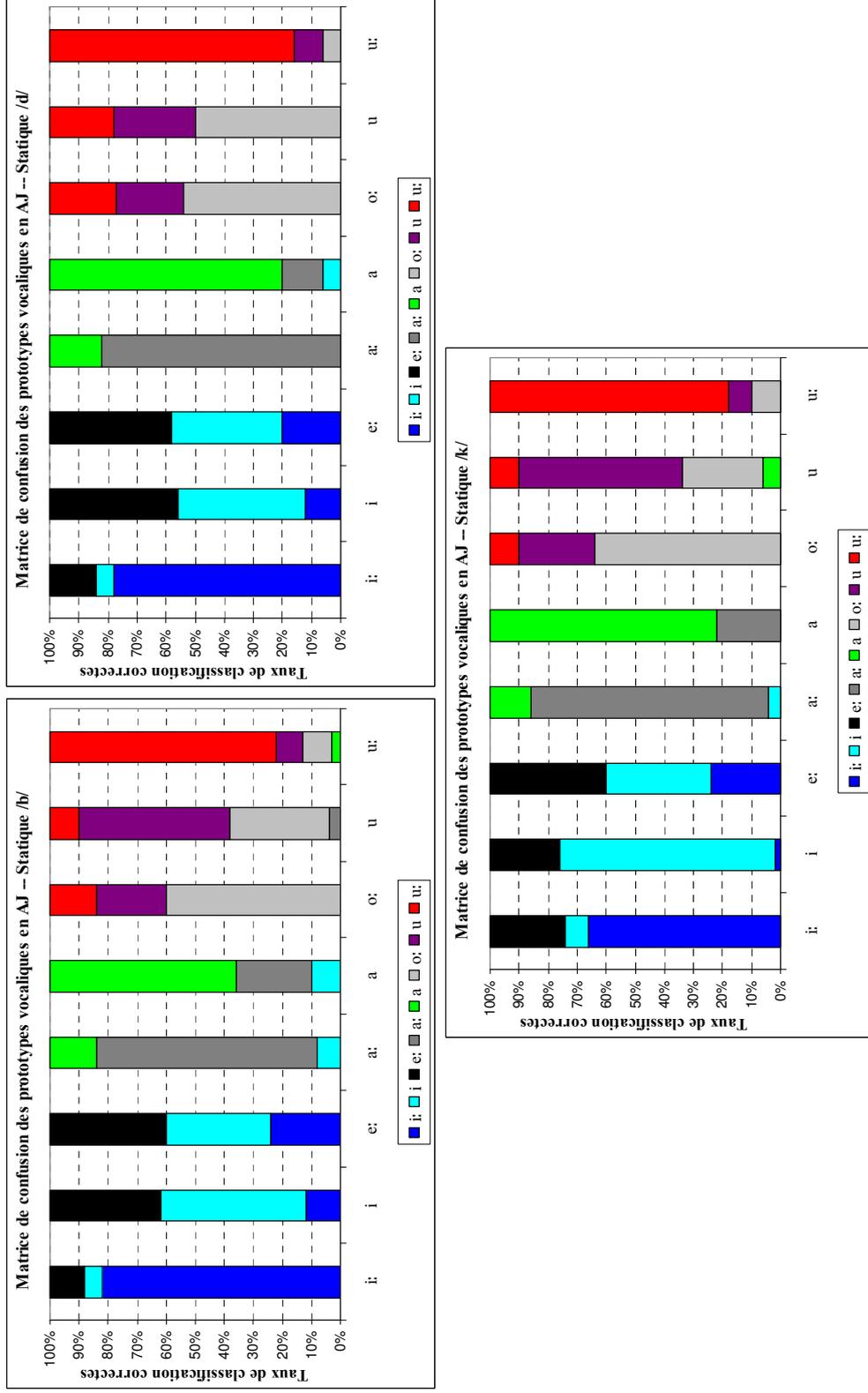


Figure 137 : Taux de classification des prototypes en AJ en statique, en fonction des contextes de présentation.

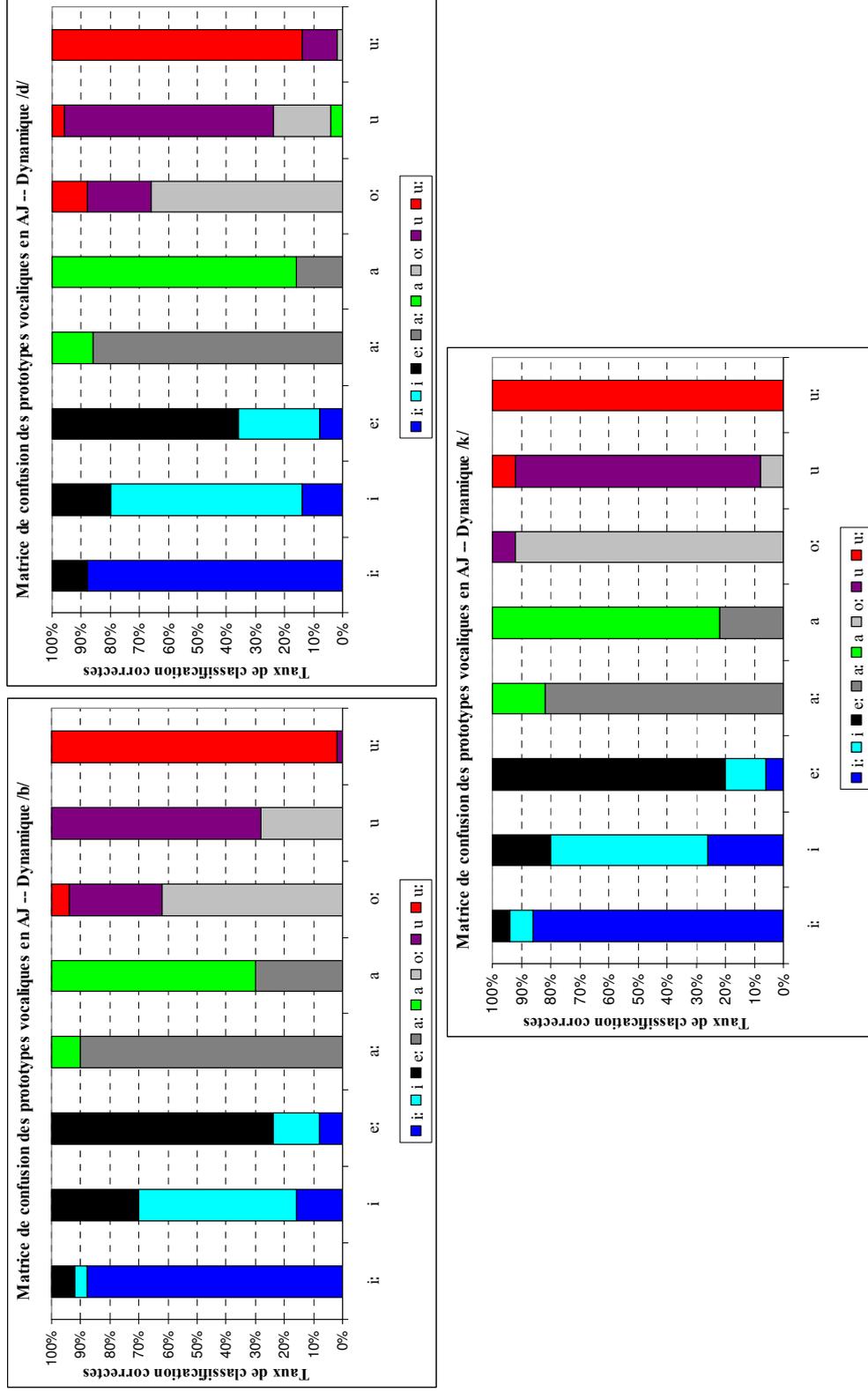


Figure 138 : Taux de classification des prototypes en AJ en dynamique, en fonction des contextes de présentation.

Les résultats d'identification des **prototypes** vocaliques obtenus en AJ montrent que l'utilisation de la **dynamique** a permis aux auditeurs jordaniens de mieux discriminer les voyelles de leur système. Les résultats de l'analyse discriminante ont permis de rendre compte de cette facilité, car le taux moyen de classification des voyelles en **dynamique** a significativement augmenté de 15% en comparaison avec les traits **statiques** (BinomialQ ; $p < 0,001$). Les auditeurs jordaniens sont conscients des différences de timbres existantes entre les trois voyelles brèves /i a u/ de leurs correspondantes longues /i: a: u:/, car les **prototypes** vocaliques qu'ils ont choisis sont les suivants : [e ə o], en **statique** et en **dynamique** et quel que soit le contexte de présentation. Les résultats de l'analyse discriminante ont permis de valider ces différences, car la plupart des confusions sont entre les couples de voyelles /i e:/ ; /ə a:/ et /u o:/. Nous observons l'homogénéité des résultats obtenus en production et en perception de la parole ; i.e., les deux voyelles brèves /i u/ sont significativement proches des voyelles longues /e: o:/. Ainsi, à partir de nos résultats, nous proposons le système vocalique perceptif en AJ comme ayant les voyelles suivantes : /i: e e: ə a: o: o u:/. Nous vérifierons de ce système à partir des résultats de la seconde expérience : la catégorisation de l'espace vocalique.

Les résultats obtenus en AM et AJ montrent que d'un côté, les choix des voyelles **prototypiques** de chaque langue ont été facilités par l'utilisation des traits **dynamiques** en comparaison avec les traits **statiques**, à partir des impressions des auditeurs de chaque langue, et de l'autre que les taux de classification correcte de l'analyse discriminante augmentent uniquement en AJ en moyenne de 15% (en AM ce taux baisse de 8%). Ces résultats vont en partie dans le sens des résultats obtenus en production où la caractérisation **dynamique** des voyelles dans les deux langues améliore les taux de classification d'au moins 20%. Nous avons observé que même en FR, la caractérisation **dynamique** des voyelles augmente le taux de classification. La prochaine étape de ce travail est de vérifier l'apport des traits **dynamiques** à l'identification des **prototypes** vocaliques en FR.

5.3.4. En Français

Comme pour les deux autres langues, nous divisons cette partie en quatre sous parties : les résultats des **prototypes** vocaliques obtenus en **statique**, en **dynamique**, la comparaison entre les deux, pour terminer avec ceux de l'analyse discriminante.

5.3.4.1. Résultats en statique (V Target)

Les **prototypes** vocaliques identifiés par les auditeurs français sont présentés sur la Figure 139. Comme pour les deux autres langues, les **prototypes** vocaliques identifiés par les auditeurs français ne sont globalement pas différents en fonction des contextes de présentation, car aucune différence significative n'est obtenue pour le contexte de présentation sur les deux axes ($F_1 : F(2, 1617)=0,59 ; p=0,56 ;$ et sur $F_2 : F(2, 1617)=0,71 ; p=0,49$). Les aires de dispersion des voyelles **prototypiques** (voir Tableau 45) ne sont pas différentes en fonction des contextes de présentation, $F(1, 4)=0,31 ; p=0,61$.

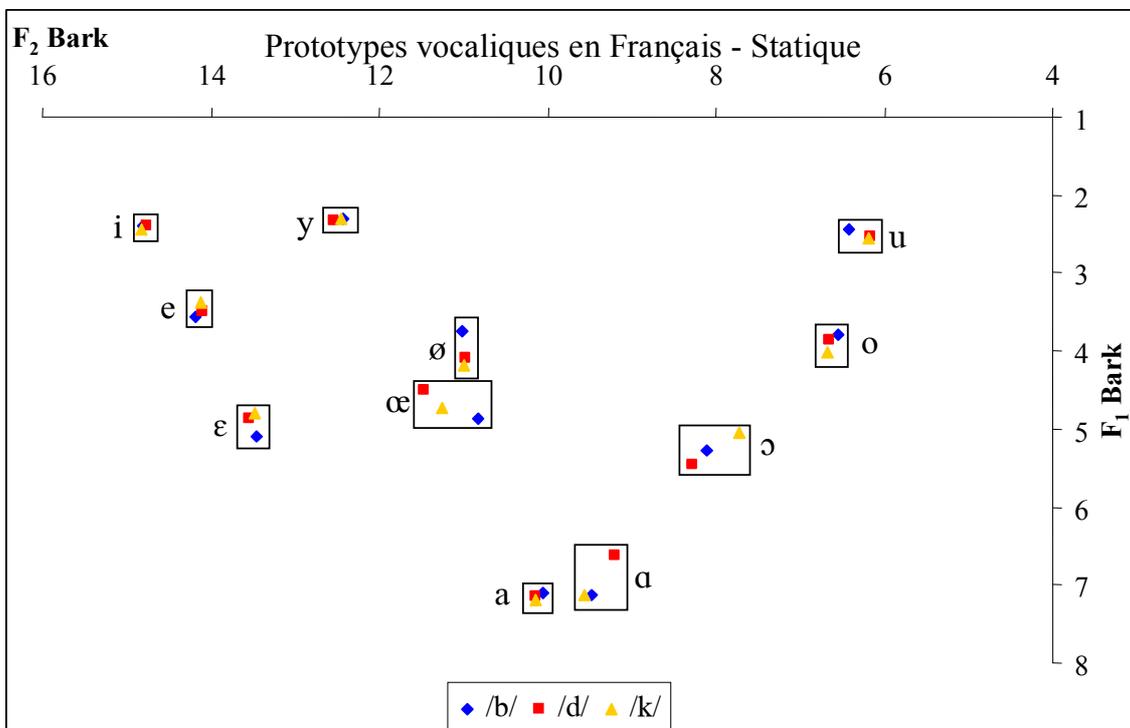


Figure 139 : Espace vocalique des **prototypes** vocaliques identifiés en FR en **statique**, en fonction des contextes de présentation.

Tableau 45 : Taille des aires de dispersion des voyelles **prototypiques** identifiées en FR en **statique** en fonction des contextes de présentation (en Bark²).

	Moyenne	Écart-Type
Bilabial	26,86	4,87
Dental	26,28	4,34
Vélaire	26,49	5,01

Comme en AM et en AJ, les espaces vocaliques obtenus en FR sont significativement non variables, i.e., stables quel que soit le contexte de présentation, conformément aux propositions de la littérature (Meunier, 2005). Toutefois, l'interaction entre les contextes de présentation et les voyelles s'est révélée significative i.e., pour une même voyelle, les choix **prototypiques** sont différents selon les contextes de présentation : pour certaines voyelles comme /œ a o ɔ/, c'est en contexte dental qu'elles sont « centralisées » sur F₁ et/ou sur F₂ ; tandis que pour d'autres, comme /y a u/, c'est le contexte bilabial qui les « centralise » (effet observé sur F₁, $F(20, 1617)=3,76$; $p<0,001$, et sur F₂, $F(20, 1617)=3,93$; $p<0,001$).

Bien que les **prototypes** vocaliques identifiés en FR soient significativement différents sur les deux axes (sur F₁, $F(10, 1617)=1379,93$; $p<0,001$; et sur F₂, $F(10, 1617)=3672,12$; $p<0,001$), nous observons un rapprochement entre les **prototypes** choisis pour le couple des voyelles /ø/ et /œ/ sur l'axe F₁, d'un côté et /a/ et /ɑ/, sur les deux axes, de l'autre. Pour ce qui est des deux autres paires de voyelles intermédiaires (/e ε/ et /o ɔ/), nous obtenons deux **prototypes** vocaliques différents pour chaque paire. Ces résultats vont en partie dans le sens des résultats obtenus en production des voyelles. Néanmoins, nous observons l'existence de différences individuelles dans le choix des **prototypes** vocaliques des voyelles médianes, visibles sur F₁, $F(9, 1617) = 116,32$; $p<0,001$; et sur F₂, $F(9, 1617) = 36,90$; $p<0,001$. L'interaction entre les auditeurs et les voyelles est significative, sur F₁, $F(90, 1617) = 19,63$; $p<0,001$, et sur F₂, $F(90, 1617) = 12,45$; $p<0,001$ (voir Figure 140, pour les résultats individuels des **prototypes** en FR).

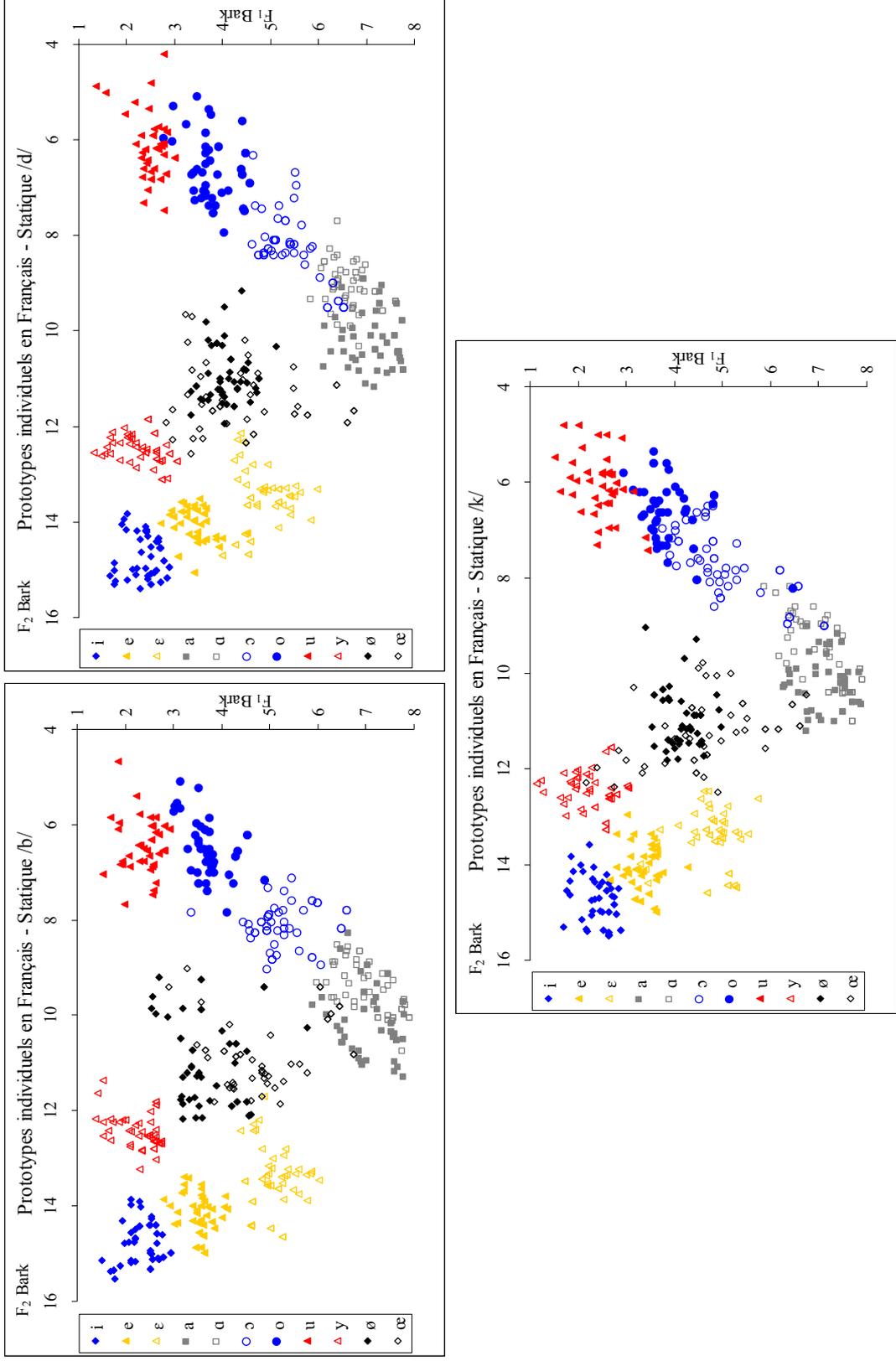


Figure 140 : Dispersion des voyelles **prototypiques** individuelles en FR en **statique**, en fonction des contextes de présentation.

Ainsi, nous présentons sur la Figure 141 la taille des ellipses de chaque voyelle **prototypique** en fonction des contextes de présentation. Nous n'obtenons aucun effet systématique pour les contextes de présentation. En revanche, nous observons que les quatre voyelles /i e y u/ obtiennent les ellipses de dispersion les plus petites. Les voyelles « médianes » présentent des ellipses très grandes.

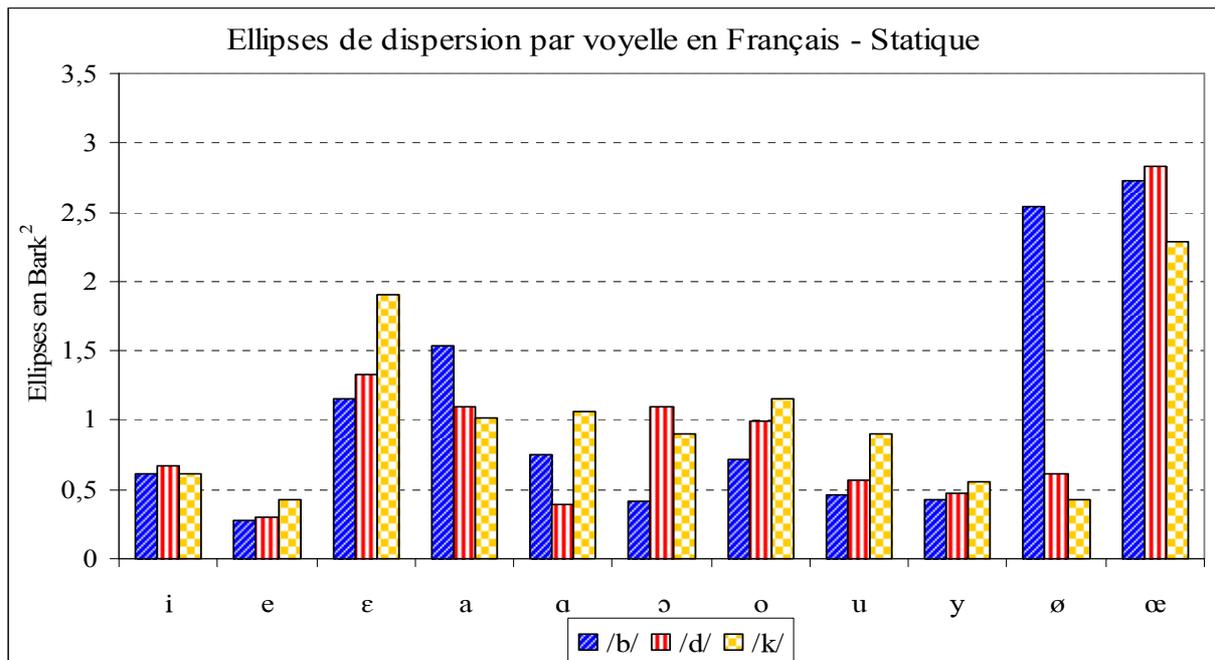


Figure 141 : Ellipses de dispersion par voyelle **prototypique** perçue en FR en **statique**, en fonction des contextes de présentation.

Les résultats que nous venons de présenter montrent que les deux voyelles ouvertes /a α/ ne sont pas significativement différentes en perception, i.e., les auditeurs français choisissent des voyelles **prototypiques** identiques pour les deux voyelles et ce quels que soient les contextes de présentation et les auditeurs. La taille des ellipses correspondant à ces deux voyelles est grande, indiquant de grands chevauchements entre les réponses des auditeurs français. Une variabilité très marquée est visible pour les voyelles « médianes » et plus particulièrement pour les deux voyelles /ø œ/, par conséquent, les résultats montrent que ces deux voyelles fusionnent dans les trois contextes de présentation : les auditeurs français choisissent des **prototypes** pour /œ/ qui varient de la voyelle [ø] jusqu'à la voyelle [a].

5.3.4.2. Résultats en dynamique (CV Target)

Les **prototypes** vocaliques moyens identifiés par les auditeurs français sont présentés sur la Figure 142. Les résultats montrent des différences significatives sur les deux axes liées au contexte de présentation : sur F_1 , $F(2, 1617) = 4,49$; $p < 0,05$, et sur F_2 , $F(2, 1617) = 10,31$; $p < 0,001$. L'interaction entre le contexte de présentation et les voyelles **prototypiques** identifiées par les auditeurs français s'est révélée significative : pour les voyelles /i y ø œ/ c'est le contexte /b/ qui les « centralise » sur les deux axes F_1 et F_2 ; pour /i a ɔ/, le contexte /b/ les « centralise » sur F_1 uniquement ; pour /e ε o u/, le contexte /d/ les « centralise » sur F_1 uniquement ; pour /a/, le contexte /d/ la « centralise » sur les deux axes F_1 et F_2 ; (sur F_1 : $F(20, 1617) = 30,96$; $p < 0,001$; sur F_2 : $F(20, 1617) = 6,96$; $p < 0,001$). Ces résultats vont à l'encontre de l'effet de stabilité de l'espace vocalique perceptif et la « normalisation » des stimuli acoustiques opérée par les auditeurs français.

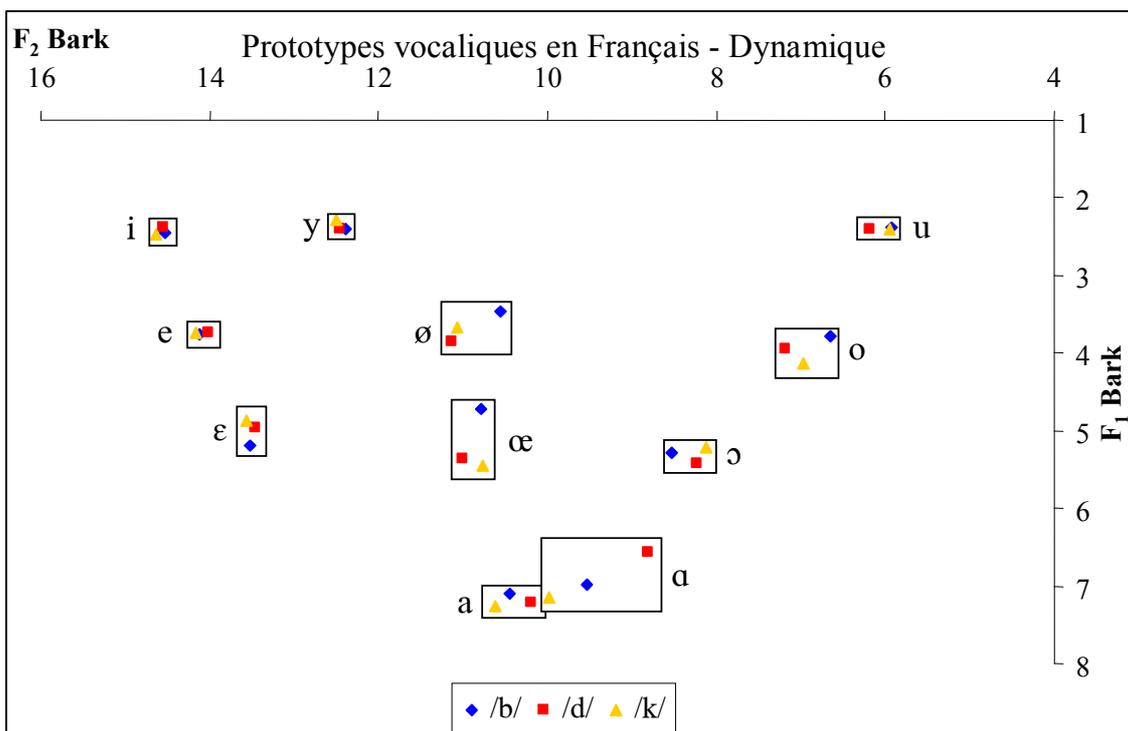


Figure 142 : Espace vocalique des **prototypes** vocaliques identifiés en FR en **dynamique**, en fonction des contextes de présentation.

Nous observons un léger effet de « périphérisation » de l'espace vocalique en contexte vélaire par rapport aux deux autres contextes. Néanmoins, il semblerait que cet effet diffère en

fonction des voyelles. En effet, nous obtenons pour les voyelles antérieures un effet d'antériorisation important en contexte vélaire par rapport aux deux autres contextes de présentation. En revanche, pour les voyelles postérieures, ce sont les deux contextes dental et vélaire qui montrent un effet de postériorisation de l'espace vocalique. La taille des aires de dispersion des voyelles en **dynamique** semble approximativement identique (avec néanmoins une différence moyenne non significative de ≈ 1 Bark² entre les contextes dental et vélaire) (voir Tableau 46).

Tableau 46 : Taille des aires de dispersion des voyelles **prototypiques** identifiées en FR en **dynamique**, en fonction des contextes de présentation (en Bark²).

	Moyenne	Écart-Type
Bilabial	26,71	4,23
Dental	26,31	3,42
Vélaire	27,08	3,83

Il existe également des différences significatives liées aux timbres de voyelles **prototypiques** choisies par les auditeurs français en **dynamique** (sur F_1 : $F(10, 1617) = 1497,83$; $p < 0,001$; et sur F_2 : $F(10, 1617) = 3482,64$; $p < 0,001$). Tous les **prototypes** vocaliques identifiés par les auditeurs français sont différents (voir Figure 142). Ainsi, toutes les voyelles « médianes » sont identifiées avec des **prototypes** moyens différents ; les deux voyelles ouvertes /a a/ quant à elles, obtiennent des **prototypes** différents, mais avec un rapprochement entre le **prototype** de /a/ en contexte /k/ et celui de /a/ en contexte /d/. Bien que ces **prototypes** soient significativement très éloignés les uns des autres, nous obtenons des ellipses plus ou moins grandes pour chaque catégorie vocalique, mais sans aucun effet systématique en fonction des contextes de présentation (voir Figure 143). Comme en **statique**, les trois voyelles /i e y u/ obtiennent les ellipses les plus petites par rapport aux autres voyelles ; les deux voyelles médianes /ø œ/ présentent les ellipses les plus grandes.

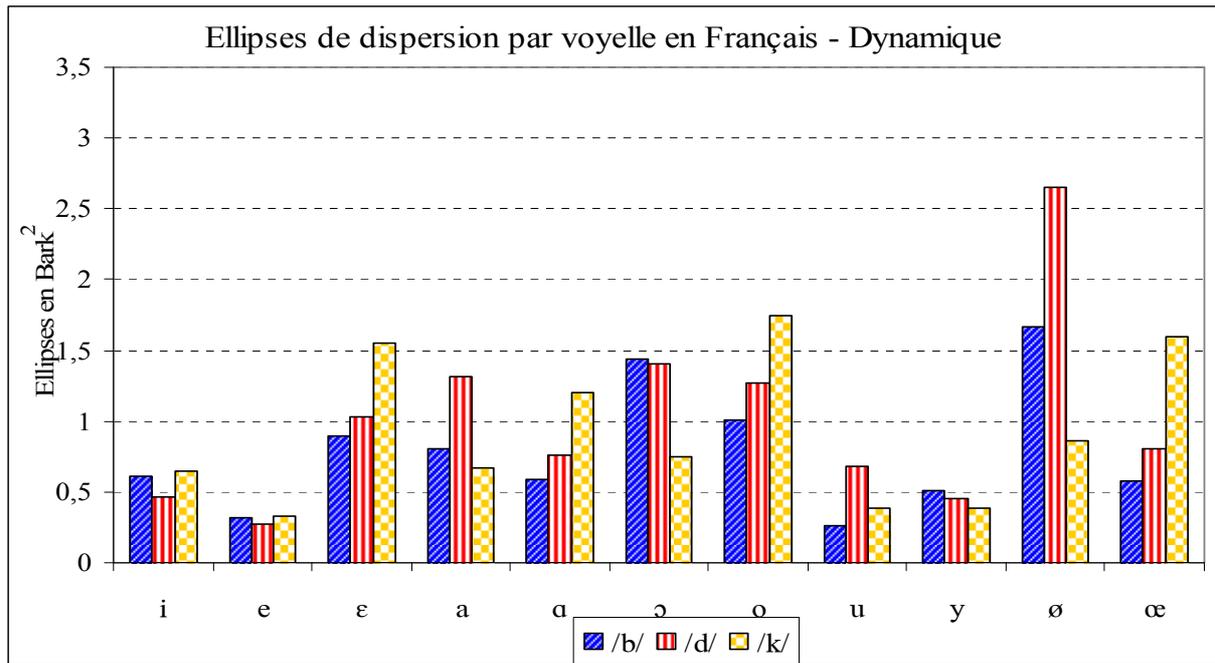


Figure 143: Ellipses de dispersion par voyelle **prototypique** perçue en FR en **dynamique**, en fonction des contextes de présentation.

Ces ellipses de taille variable indiquent l'importante variabilité inter-individuelle (voir Figure 144). En effet, les résultats statistiques montrent qu'il y a une différence significative, sur les deux axes F_1 et F_2 , liée aux auditeurs (sur F_1 , $F(9, 1617) = 83,97$; $p < 0,001$, et sur F_2 , $F(9, 1617) = 36,08$; $p < 0,001$) et à l'interaction entre les auditeurs et les voyelles (sur F_1 , $F(90, 1617) = 12,21$; $p < 0,001$; sur F_2 , $F(90, 1320) = 6,70$; $p < 0,001$). L'interaction entre les auditeurs, les voyelles et le contexte de présentation est significative (sur F_1 , $F(180, 1617) = 4,07$; $p < 0,001$; sur F_2 , $F(180, 1320) = 2,24$; $p < 0,001$).

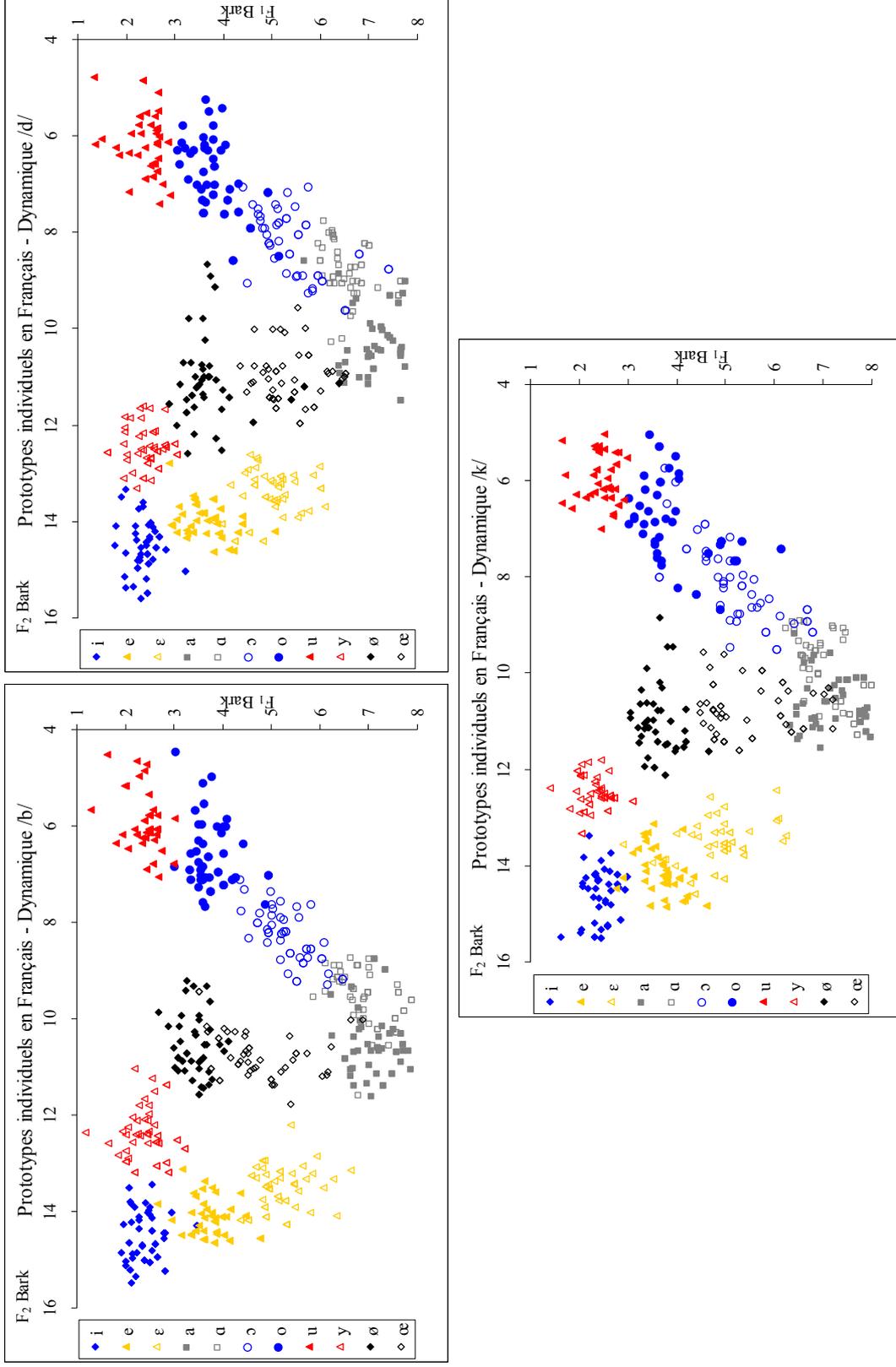


Figure 144 : Dispersion des voyelles **prototypiques** individuelles en FR en **dynamique**, en fonction des contextes de présentation.

Les graphiques présentés sur la Figure 144, montrent que malgré la variabilité importante dans le choix des **prototypes** vocaliques en FR, les auditeurs français choisissent des **prototypes** qui se distinguent les uns des autres : seules les deux voyelles /a ɑ/ sont confondues entre elles ; les deux voyelles « médianes » /ø œ/ sont significativement différentes quels que soient les contextes de présentation et les locuteurs. Nous vérifions l'apport de la **dynamique** en FR, dans la partie qui suit.

5.3.4.3. Comparaison des résultats en statique vs en dynamique

Les résultats de la comparaison **statique** vs **dynamique** en perception des voyelles **prototypiques** identifiées en FR montrent des différences significatives, uniquement sur F_1 (voir Figure 145) : l'espace vocalique en **dynamique** semble significativement plus ouvert que celui en **statique** ($F(1, 3234) = 12,82$; $p < 0,001$). Cette différence n'affecte en rien la taille de l'espace vocalique dans les deux modalités. Ainsi, l'espace vocalique perceptif en **statique** obtient une aire de 26,54 Bark² tandis qu'en **dynamique** l'aire est de 26,70 Bark² (différence non significative : $F(1, 4) = 0,31$; $p = 0,61$).

L'interaction entre les traits utilisés (**statique** vs **dynamique**) et les voyelles est significative. En **dynamique**, toutes les voyelles /e a o ɔ y œ/ sont significativement plus « ouvertes », tandis que les voyelles /u ø/ sont significativement plus « fermées », sur l'axe F_1 ($F(10, 3234) = 2,09$; $p < 0,05$) ; sur l'axe F_2 , les voyelles /a ɔ o/ sont plus « antériorisées » ; les /i u ø œ / sont significativement plus « postérieures » ($F(10, 3234) = 9,38$; $p < 0,001$). Sur la Figure 146, nous présentons la comparaison entre les **prototypes** vocaliques identifiés par les auditeurs français, en fonction des contextes de présentation, de la voyelle et des traits utilisés. L'interaction entre les traits utilisés et les contextes de présentation est significative : en **dynamique**, l'espace vocalique obtenu en contexte /k/ est plus « antérieur » ($F(2, 3234) = 6,14$; $p < 0,005$). Nous observons également des différences importantes en fonction du choix des voyelles **prototypiques** en FR qui semblent dépendre à la fois des traits utilisés et des contextes de présentation, ce sur les deux axes : sur F_1 , les voyelles identifiées en contextes /b k/ en **dynamique** semblent être plus « ouvertes » ($F(20, 3234) = 19,21$; $p < 0,001$), et sur

F₂, le contexte /d/ semble « centraliser » les voyelles en **dynamique** ($F(20, 3234) = 2,14$; $p < 0,005$), (voir Tableau 47).

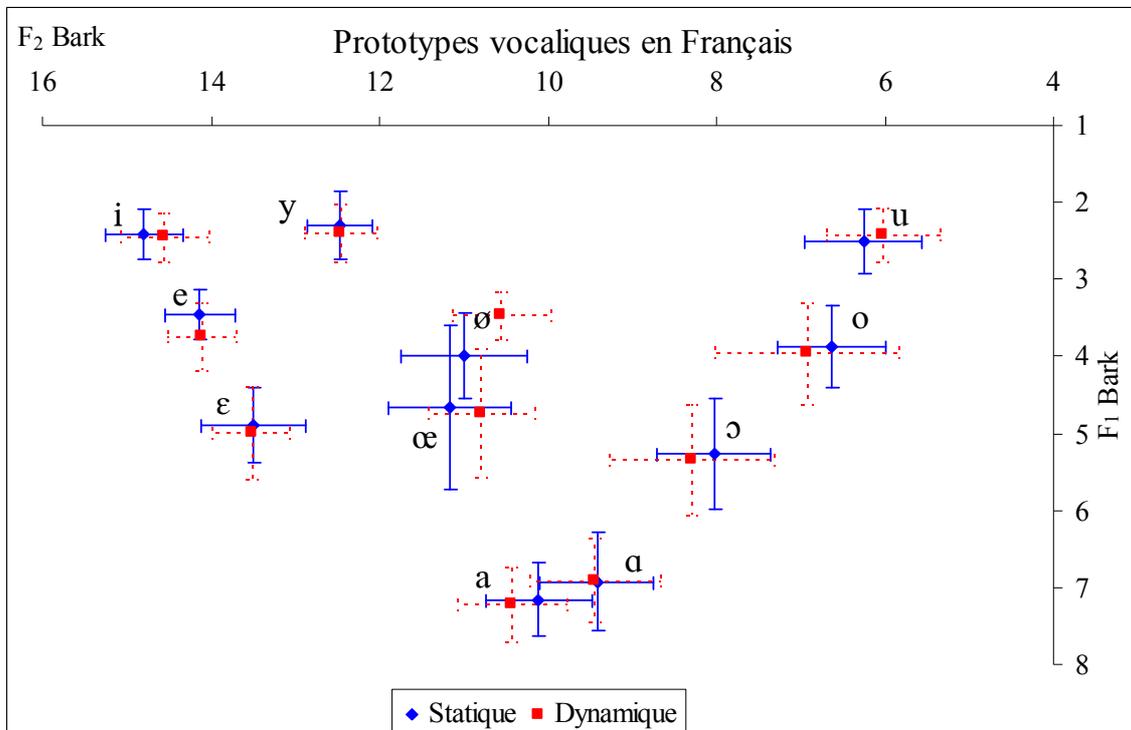


Figure 145 : Espaces vocaliques des **prototypes** FR en **statique** vs **dynamique**.

Tableau 47 : Résultats des T-test de Student entre les voyelles **prototypiques** identifiées par les auditeurs français, en **statique** vs **dynamique**. [+F] = [+fermé], [+O] = [+ouvert], [+A] = [+antérieur], [+P] = [+postérieur], NS = Non Significative. Le tableau se lit de la façon suivante : /i/ > /b/ > F₁ = [+O], $p < 0,05$: la voyelle /i/ identifiée en contexte /b/ en **dynamique** est [+ouverte], $p < 0,05$, par rapport à celle en **statique**.

	F ₁			F ₂		
	/b/	/d/	/k/	/b/	/d/	/k/
/i/	[+O], $p < 0,05$	NS	NS	[+P], $p < 0,001$	[+P], $p < 0,001$	[+P], $p < 0,005$
/e/	[+O], $p < 0,001$	[+O], $p < 0,001$	[+O], $p < 0,001$	NS	NS	NS
/ε/	[+O], $p < 0,05$	[+O], $p < 0,05$	NS	NS	NS	[+A], $p < 0,05$
/a/	NS	NS	[+O], $p < 0,05$	[+A], $p < 0,001$	NS	[+A], $p < 0,001$
/ɑ/	[+F], $p < 0,05$	NS	NS	NS	[+P], $p < 0,001$	[+A], $p < 0,001$
/ɔ/	NS	NS	NS	[+A], $p < 0,02$	NS	[+A], $p < 0,005$
/o/	NS	NS	NS	NS	[+A], $p < 0,01$	[+A], $p < 0,05$
/u/	NS	[+F], $p < 0,01$	[+F], $p < 0,05$	[+P], $p < 0,001$	NS	[+P], $p < 0,05$
/y/	[+O], $p < 0,05$	[+O], $p < 0,05$	NS	NS	NS	NS
/ø/	[+F], $p < 0,005$	[+F], $p < 0,05$	[+F], $p < 0,001$	[+P], $p < 0,005$	[+A], $p < 0,05$	NS
/œ/	NS	[+O], $p < 0,001$	[+O], $p < 0,001$	NS	[+P], $p < 0,001$	[+P], $p < 0,001$

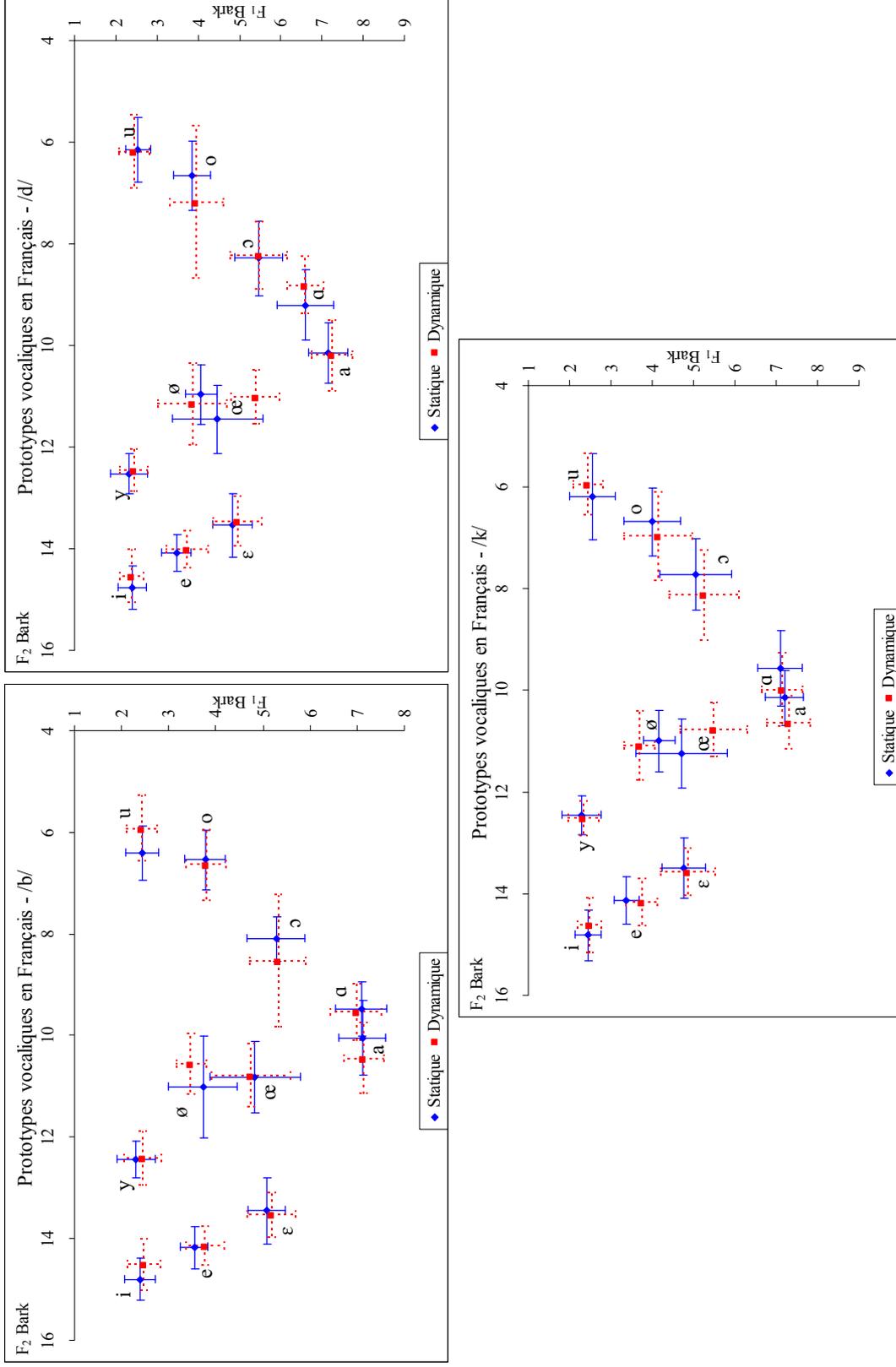


Figure 146 : Espace vocalique des prototypes en FR en statique vs dynamique, en fonction des contextes de présentation.

La plupart des interactions se sont révélées significatives. Néanmoins, nous n’obtenons aucune différence significative en fonction de la taille des aires de dispersion des voyelles en fonction des contextes de présentation (voir Figure 147) : $F(1, 4) = 0,31$; $p = 0,61$. Les auditeurs français choisissent des voyelles significativement différentes, mais sans aucune modification de la taille de l’espace perceptif.

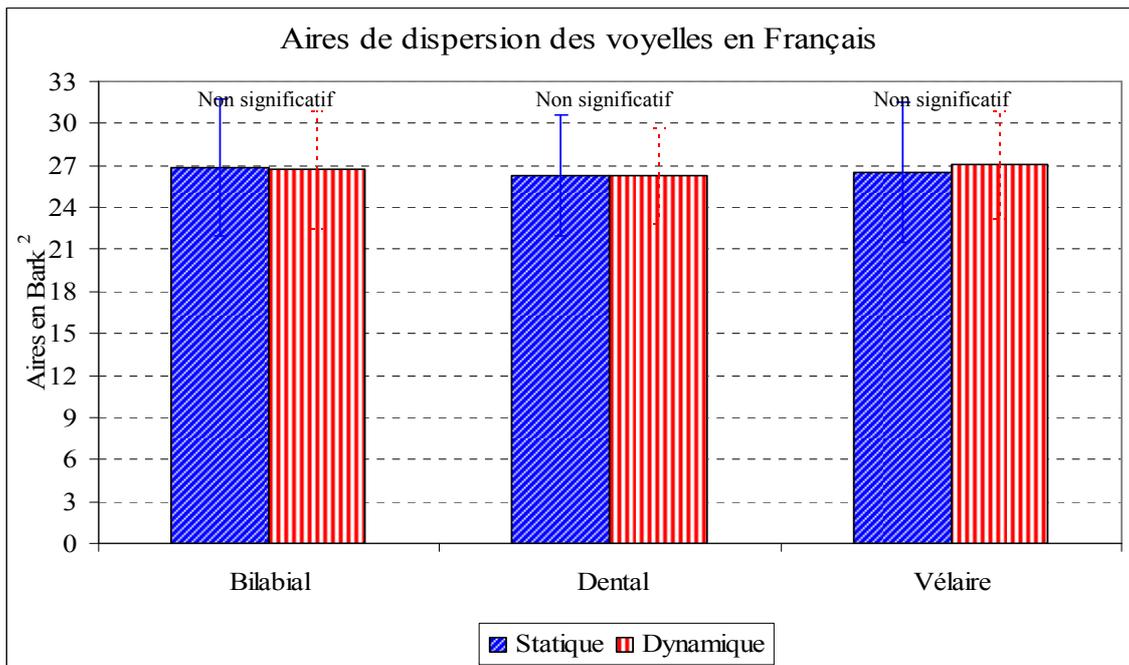


Figure 147 : Aires de dispersion des voyelles **prototypiques** identifiées en FR en **statique** vs **dynamique**, en fonction des contextes de présentation.

Nous avons voulu comparer la taille des ellipses de chaque catégorie vocalique perçue en **statique** vs **dynamique** afin de voir l’incidence des traits **dynamiques** dans la précision du choix des auditeurs français. Ainsi, nous présentons sur la Figure 148 les résultats de la taille des ellipses par voyelle **prototypique** en fonction des traits utilisés. L’utilisation de la **dynamique** donne plus de précision aux auditeurs français dans le choix des **prototypes** de leur système, surtout pour les voyelles /i e ε a u y ø œ/, tandis que pour les voyelles /ɑ ɔ o/, la **dynamique** provoque des variations inter-individuelles importantes dans le choix des **prototypes** (effet non significatif, pour les résultats en fonction des contextes de présentation, voir Annexe 35).

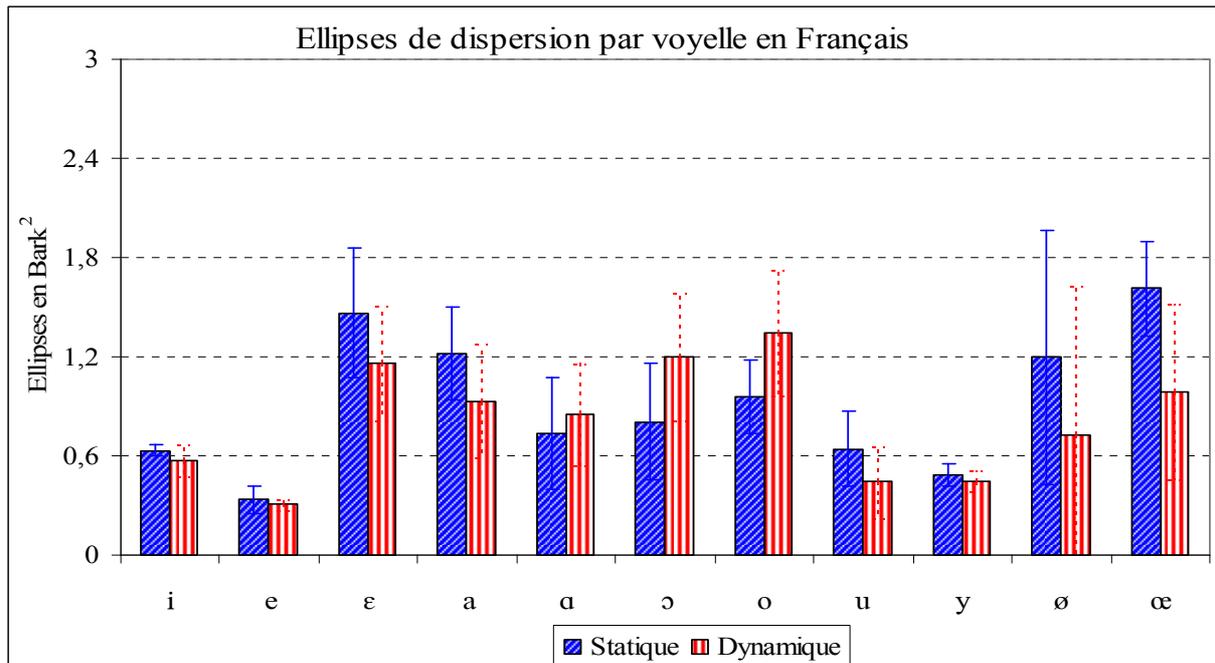


Figure 148 : Ellipses de dispersion par voyelle **prototypique** perçue en FR en **statique** vs **dynamique**.

Les résultats du questionnaire donné aux auditeurs français montrent que pour ces derniers, la tâche d'identification des voyelles **prototypiques** en **statique** était beaucoup plus difficile que celle en **dynamique**. Les auditeurs ont expliqué que les sons de synthèse paraissaient « *bizarres* », néanmoins ils n'ont éprouvé aucune difficulté lors de l'identification des voyelles avec les deux modalités. Ils ont expliqué qu'en **dynamique**, ils trouvaient « *la direction à suivre* » pour identifier les **prototypes**, c'est pourquoi nous observons des positions différentes des **prototypes** en fonction des contextes de présentation. Nous tenterons de vérifier la validité des résultats obtenus avec l'aide de l'analyse discriminante qui nous donnera plus d'explication quant au rôle de la **dynamique** dans l'identification des **prototypes** du FR.

5.3.4.4. Résultats de l'analyse discriminante

Les paramètres utilisés pour cette analyse sont les valeurs de F_1 et de F_2 des **prototypes** vocaliques identifiés par les auditeurs français.

Nos résultats ont montré que les **prototypes** vocaliques identifiés par les auditeurs français sont significativement différents quels que soient les traits utilisés, les contextes de

présentation et les voyelles. Les résultats de l'analyse discriminante confirment cela. Ainsi, nous présentons dans le Tableau 48 les différents taux de classification des **prototypes** en **statique** vs **dynamique** et par contexte de présentation.

Tableau 48 : Taux de classification correcte des voyelles **prototypiques** identifiées en FR en **statique** vs **dynamique** en fonction des contextes de présentation. (Taux significatifs, Box's M ; $p < 0,001$).

	/b/	/d/	/k/
Statique	85,1%	82,5%	78,5%
Dynamique	86,5%	85,8%	80,4%

Le taux moyen de classification de l'analyse discriminante obtenu en FR est significativement plus élevé de 2,07% en **dynamique** qu'en **statique** (BinomialQ ; $p < 0,001$). Les taux moyens par modalité sont relativement élevés : 82,03% en **statique** contre 84,1% en **dynamique**. Ceci montre que les auditeurs français sont capables d'identifier « correctement » les voyelles de leur système à partir de cette expérience de perception. Nous examinons ainsi les classifications correctes et les confusions obtenues à partir de l'analyse discriminante.

En **statique** et dans les trois contextes de présentation (voir Figure 149), nous observons que 98% à 100% des voyelles fermées /i y u/ ont été classifiées correctement. Les quatre voyelles « médianes » /e ε o ɔ/ ont été classifiées avec des taux élevés (de 74% à 100%). La plupart des confusions concernaient les couples de voyelles suivantes : /a α/, /a ɔ/ et /ø œ/. Ces confusions sont conformes à nos résultats sur la taille des ellipses (voir Figure 141) et la dispersion individuelle des voyelles (voir Figure 140).

En **dynamique** en revanche, nous observons que les trois voyelles fermées /i y u/ obtiennent des taux très élevés proches des 100% lors de leur classification ; les voyelles « médianes » /e ε ø œ o ɔ/ obtiennent des taux intermédiaires et les deux voyelles ouvertes /a α/ sont confondues soit entre elles soit avec /ɔ/ (voir Figure 150). Bien que le taux moyen de classification correcte en **dynamique** soit significativement plus élevé qu'en **statique**, nous observons plus de confusions, d'une façon générale, dans la classification des **prototypes** en **dynamique**.

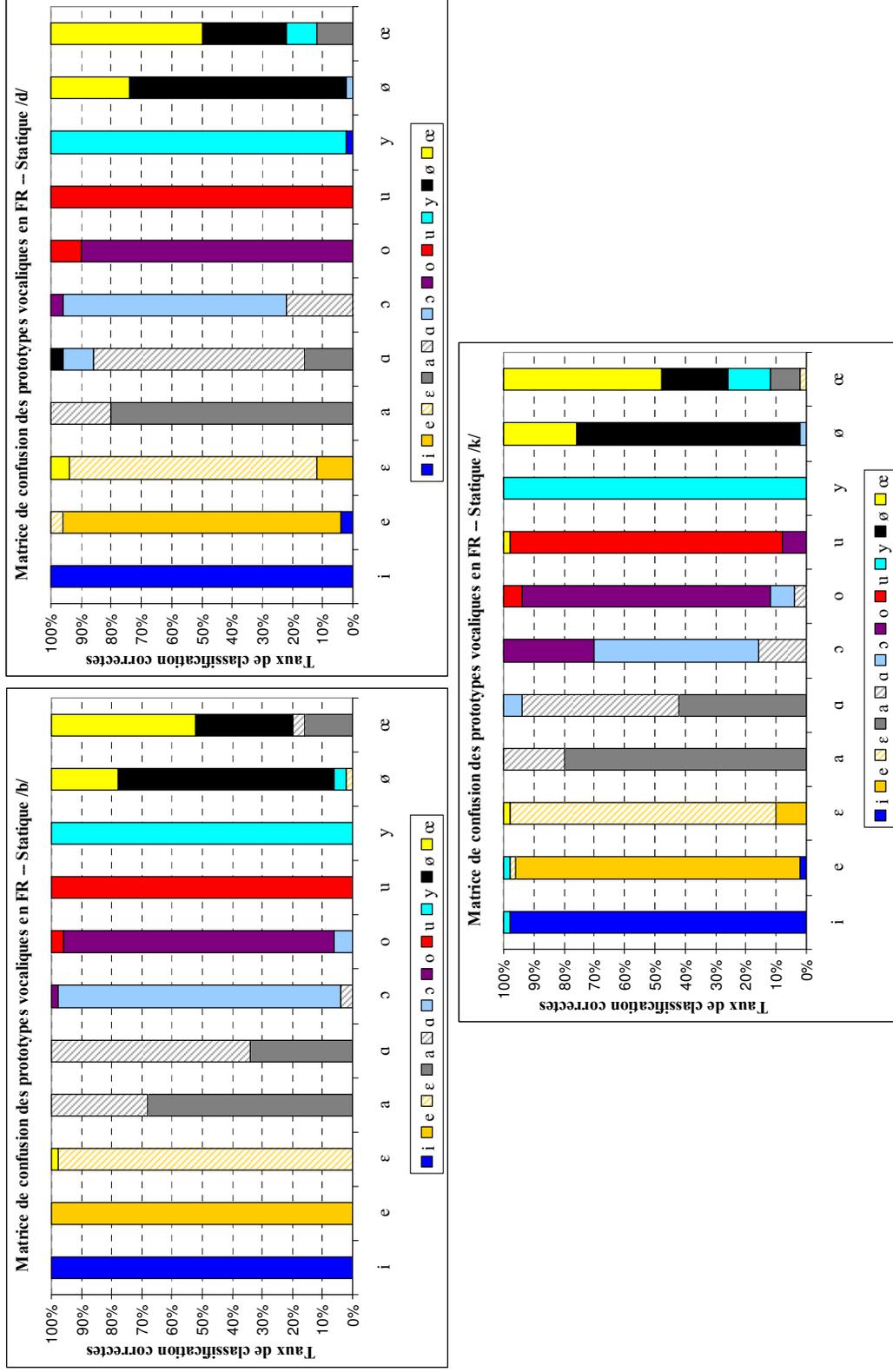


Figure 149 : Taux de classification des prototypes en FR en statique, en fonction des contextes de présentation.

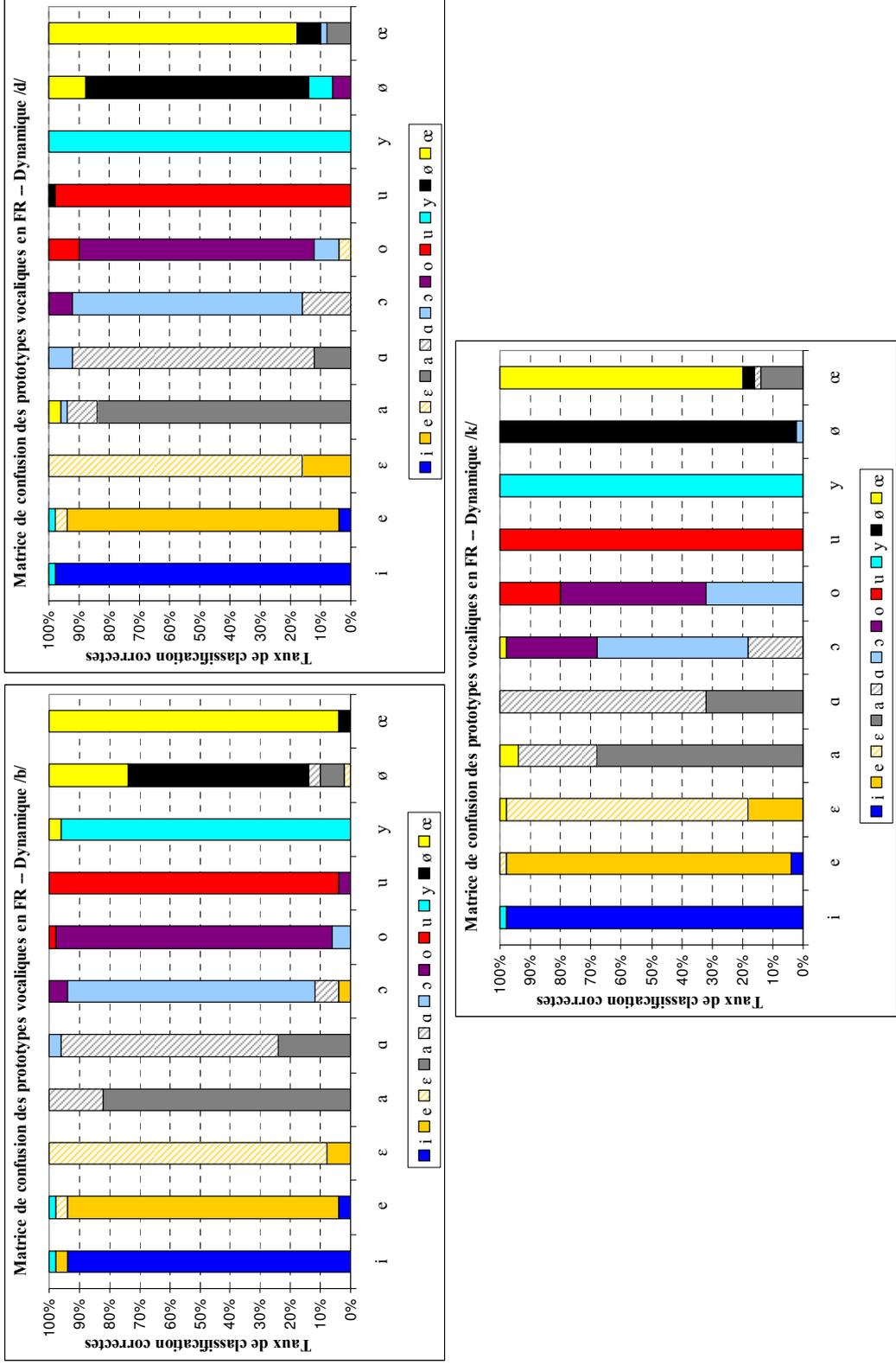


Figure 150 : Taux de classification des prototypes en FR en dynamique, en fonction des contextes de présentation.

Les **prototypes** vocaliques identifiés par les auditeurs français en **statique** vs **dynamique**, sont significativement différents. Le contexte de présentation a affecté les **prototypes** vocaliques choisis par les auditeurs français uniquement en **dynamique**. Les résultats de l'analyse discriminante ont montré la possibilité de discriminer les voyelles en français en fonction des traits utilisés et des contextes de présentation avec des taux moyens élevés. Nous avons observé l'amélioration significative, avec un taux moyen de 2%, de la classification des voyelles du FR par l'analyse discriminante. Tous les auditeurs français ont trouvé la tâche d'identification des **prototypes** en **statique** plus difficile que celle en **dynamique** ; cette dernière a contribué à la meilleure identification des **prototypes** de leur système. Il semblerait que la **dynamique** soit un paramètre très important dans l'identification des voyelles dans les trois langues étudiées.

Nous venons d'examiner les résultats des expériences de perception des voyelles proposées aux auditeurs des trois langues étudiées. En production de la parole, nous avons observé l'existence de différences dans l'organisation des systèmes vocaliques dans les trois langues, liée à la densité des systèmes. Nous voulons examiner si de telles différences peuvent exister (ou non) en perception. On sait que la perception de la parole, et plus particulièrement des voyelles, propose que l'organisation des espaces vocaliques perceptifs dépende de la faculté de reconnaissance des variations dans le signal acoustique et de la normalisation de celles-ci. Ainsi, la littérature propose que, comme la perception de la parole est un phénomène *many-to-one*, la normalisation que le signal acoustique subi pour permettre l'identification des « représentations mentales » des voyelles, dans notre cas, affecte la dispersion des voyelles dans cet espace perceptif. L'espace vocalique perceptif est un espace homogène entre tous les êtres humains, on acquiert notre langue maternelle et modifie notre perception des sons contenus dans notre langue et dans les langues que l'on connaît, mais cet espace n'est pas changeable. Ainsi, on peut prétendre que les différences observables en production de la parole ne pourraient exister en perception, i.e., l'espace vocalique perceptif doit être homogène pour toutes les langues (sa taille !!) mais c'est au niveau de l'organisation interne de l'espace que des différences peuvent apparaître (voir les travaux de Kuhl, 1991 ; Kuhl & Iverson, 1995* ; Iverson & Kuhl, 1996 ; Meunier, 2005 ; etc...).

Nous nous intéressons dans la partie qui suit aux résultats comparatifs inter-langues. Ainsi, nous vérifierons si la densité des systèmes vocaliques peut affecter la taille de l'espace vocalique perceptif global et la dispersion des voyelles à l'intérieur de cet espace et plus spécifiquement la position des trois voyelles **prototypiques** extrêmes /i a u/ (voyelles longues en arabe). Nous avons choisi d'effectuer cette étude comparative en fonction des trois voyelles /i a u/ car ce sont les trois seules voyelles communes entre les trois langues étudiées et elles délimitent l'espace vocalique. Au besoin, nous nous servirons des résultats obtenus pour chaque langue (qui sont présentés dans les parties précédentes) afin de renforcer nos résultats.

5.3.5. Résultats Inter-Langues

Nous commençons cette partie par la présentation des résultats en **statique**, ensuite en **dynamique**. Nous poursuivons par la comparaison des résultats dans les deux modalités et terminons par les résultats de l'analyse discriminante.

5.3.5.1. Résultats en statique (V Target)

La comparaison entre les espaces vocaliques perceptifs obtenus à partir des **prototypes** de /i a u/ identifiés par les auditeurs des trois langues est présentée graphiquement sur la Figure 151. L'observation des trois espaces vocaliques montre des différences significatives en fonction des langues, visibles sur les deux axes : l'espace vocalique perceptif est significativement plus « périphérique » en FR qu'en AJ ; c'est l'AM qui présente l'espace le plus « centralisé » (sur F_1 : $F(2, 1321)=22,14$; $p<0,001$; sur F_2 : $F(2, 1321)=40,30$; $p<0,001$). Les trois voyelles **prototypiques** /i a u/ sont positionnées différemment entre les trois langues. Ainsi, sur F_1 , la voyelle /a/ est significativement plus « ouverte » en FR et plus « fermée » en AM qu'en AJ et les deux voyelles /i u/ sont significativement plus « fermées » en FR et plus « ouvertes » en AM qu'en AJ ($F(4, 1321)=98,31$; $p<0,001$). Sur F_2 , la voyelle /a/ est significativement plus « antérieure » en AJ et plus « postérieure » en FR ; /i/ est plus antérieure en FR qu'en AM ou en AJ et /u/ est plus postérieure en FR qu'en AM ou en AJ ($F(4, 1321)=67,72$; $p<0,001$).

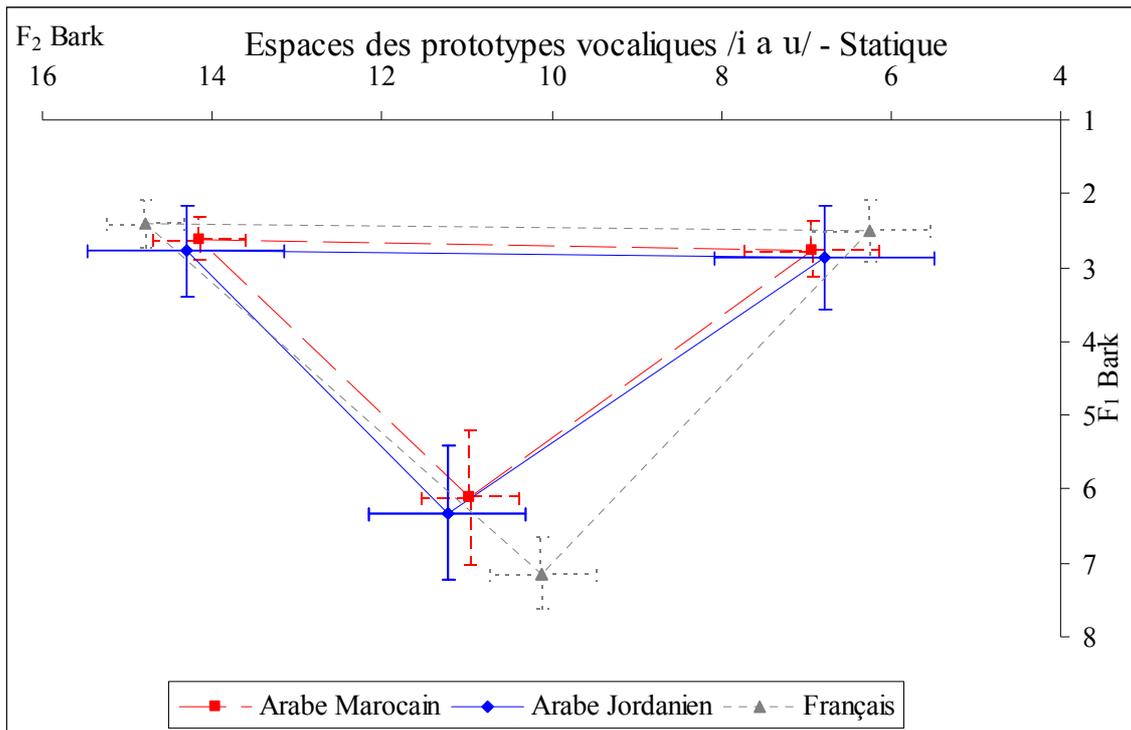


Figure 151 : Espaces des **prototypes** vocaliques /i a u/ dans les trois langues en **statique**.

La « centralisation » de l'espace vocalique perceptif en AM par rapport aux deux autres langues et plus particulièrement le FR est confirmée quel que soit le contexte de présentation (voir Figure 152) : aucun changement significatif n'est détecté en fonction des interactions entre les langues et les contextes de présentation d'un côté, ou entre les langues, les contextes de présentation et les voyelles, de l'autre. Autrement dit, ces résultats sont cohérents quels que soient les contextes de présentation et les voyelles. Nous avons voulu savoir si la « centralisation » de l'espace en AM et la « périphérisation » en FR affectent la taille de l'espace vocalique perceptif obtenu. Ainsi, les aires de dispersion obtenues à partir des trois voyelles **prototypiques** /i a u/ sont statistiquement non différentes entre l'AM et l'AJ (12,49 Bark² en AM, 13,79 Bark² en AJ : $F(1, 58)=1,24$; $p=0,27$), en revanche, elles le sont entre d'un côté, le FR et l'AM (12,49 Bark² en AM et 20,12 Bark² en FR : $F(1, 58)=93,51$; $p<0,001$), et de l'autre, le FR et l'AJ (13,79 Bark² en AJ et 20,12 Bark² en FR : $F(1, 58)=30,65$; $p<0,001$).

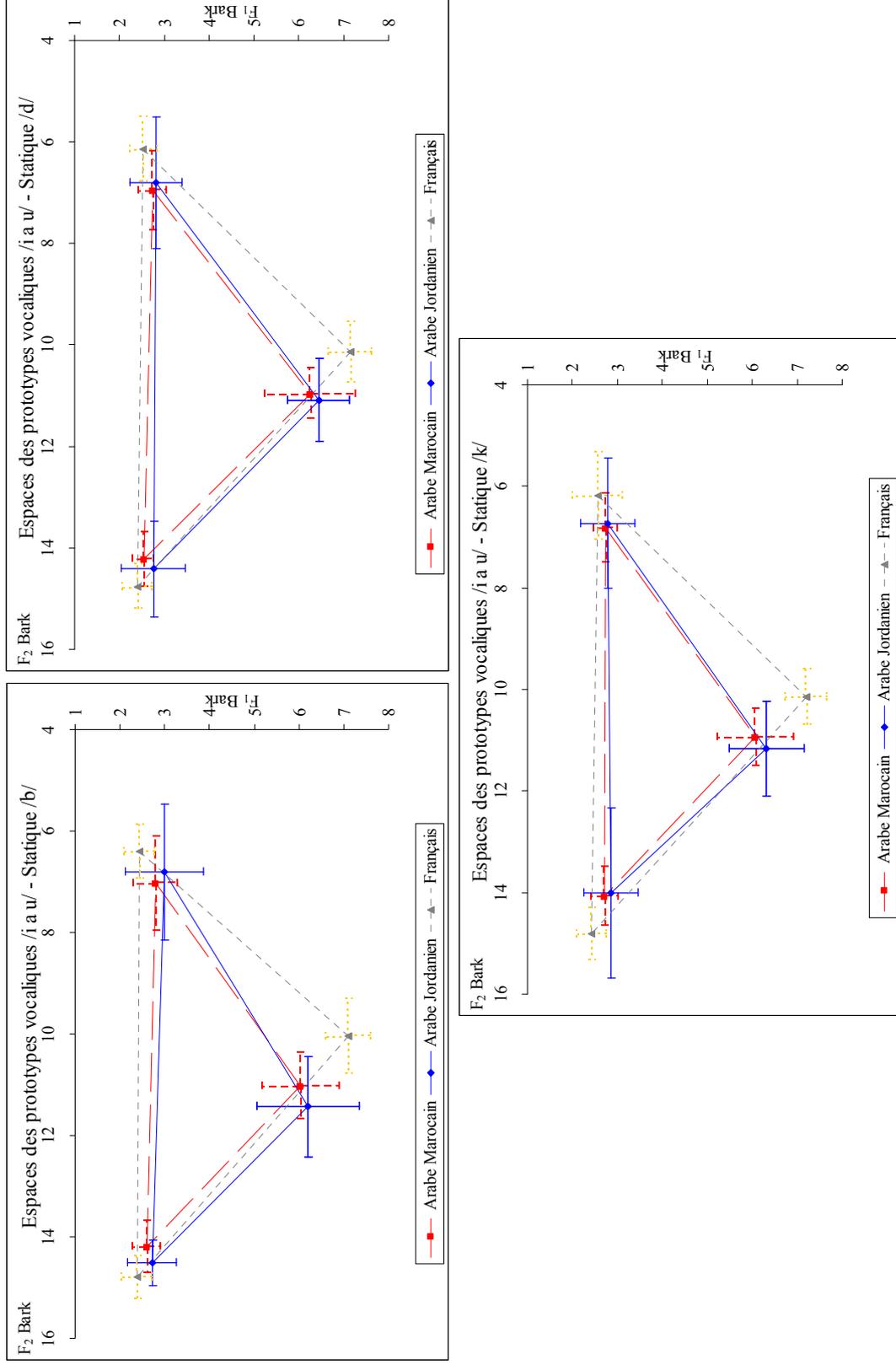


Figure 152 : Espaces des **prototypes** vocaliques /i a u/ dans les trois langues en **statique**, par contexte de présentation.

L'effet de la densité des systèmes vocaliques semble affecté la taille de l'espace obtenu en FR seulement, car aucune différence n'est obtenue entre l'AM et l'AJ. La taille des aires de dispersion plus grande obtenue en FR par rapport à l'AM et l'AJ va dans le sens des propositions de la **théorie de la dispersion adaptée**, qui explique que la densité des systèmes affecte l'organisation des espaces vocaliques : plus une langue a des voyelles et plus la taille des espaces vocaliques est grande, afin de maintenir la distinction entre les différentes voyelles. En revanche, la taille des aires non différente obtenue en AM et en AJ conforte les propositions de la **théorie quantique**, qui expliquent que la densité des systèmes n'affecte pas la taille des espaces vocaliques.

Cette « centralisation » de l'espace vocalique perceptif en AM et en AJ comparée à celui du FR est fonction du contexte de présentation : effet visible sur la taille des aires de dispersion des espaces vocaliques (voir Figure 153). Les aires de dispersion obtenues en FR sont significativement plus grandes que celles en AM ou en AJ ; aucune différence n'est obtenue pour les aires de dispersion des espaces vocaliques des deux dialectes arabes. Encore une fois, ce résultat va à l'encontre de l'hypothèse de stabilité des espaces vocaliques perceptifs quelles que soient la langue et la densité des systèmes. Il semblerait donc que la densité des systèmes affecte bel et bien la taille des aires de dispersion des espaces vocaliques perceptifs. Les aires de dispersion des voyelles **prototypiques** identifiées sont plus petites en contexte bilabial, plus grandes en contexte dental (pour AM et AJ seulement) et intermédiaires en contexte vélaire (sauf en FR où la différence est la suivante : Bilabial < Dental ≤ Vélaire).

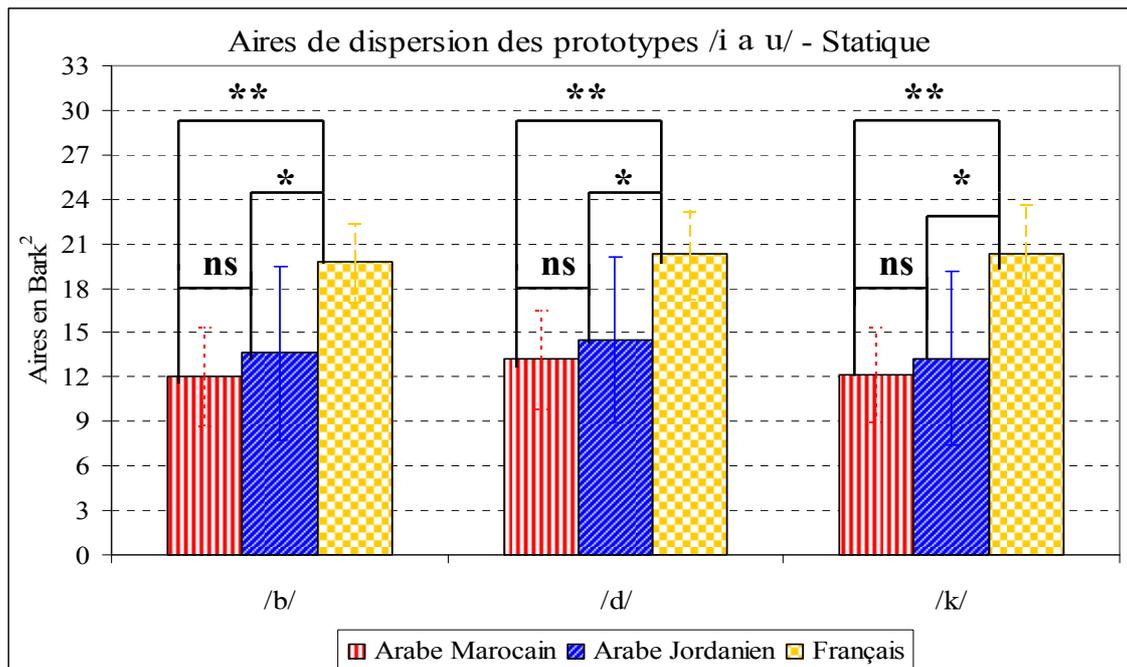


Figure 153 : Aires de dispersion des voyelles **prototypiques** /i a u/ dans les trois langues en **statique**, en fonction du contexte de présentation. (ns = non significative ; * = $p < 0,01$; ** = $p < 0,001$).

Les résultats que nous venons de présenter sont partiellement en accord avec l'hypothèse selon laquelle les espaces vocaliques perceptifs dans les trois langues dépendent de la densité de leurs systèmes.

Nous voulons ensuite voir si l'effet de la densité des systèmes vocaliques peut affecter la taille des ellipses de chaque voyelle **prototypique**. Ainsi, nous présentons sur la Figure 154 la taille des ellipses de dispersion par voyelle **prototypique** en fonction des trois langues. Les résultats montrent que c'est en AJ que les ellipses sont les plus grandes ; l'AM présente les ellipses les plus petites pour les deux voyelles /i u/ tandis que le FR se situe entre les deux. Les explications apportées par Patricia Kuhl à partir du modèle « *The Native Language Magnet* » (Kuhl & Iverson, 1995*, voir § 2.3.2.6) vont dans le sens d'une modification des positions des voyelles à l'intérieur de l'espace vocalique dans les langues du monde, mais ne prédit en rien les différences de taille des ellipses vocaliques. En effet, ces résultats vont à l'encontre des propositions de stabilité des systèmes vocaliques, mais mettent en évidence le rôle de la densité des systèmes dans l'organisation des espaces vocaliques.

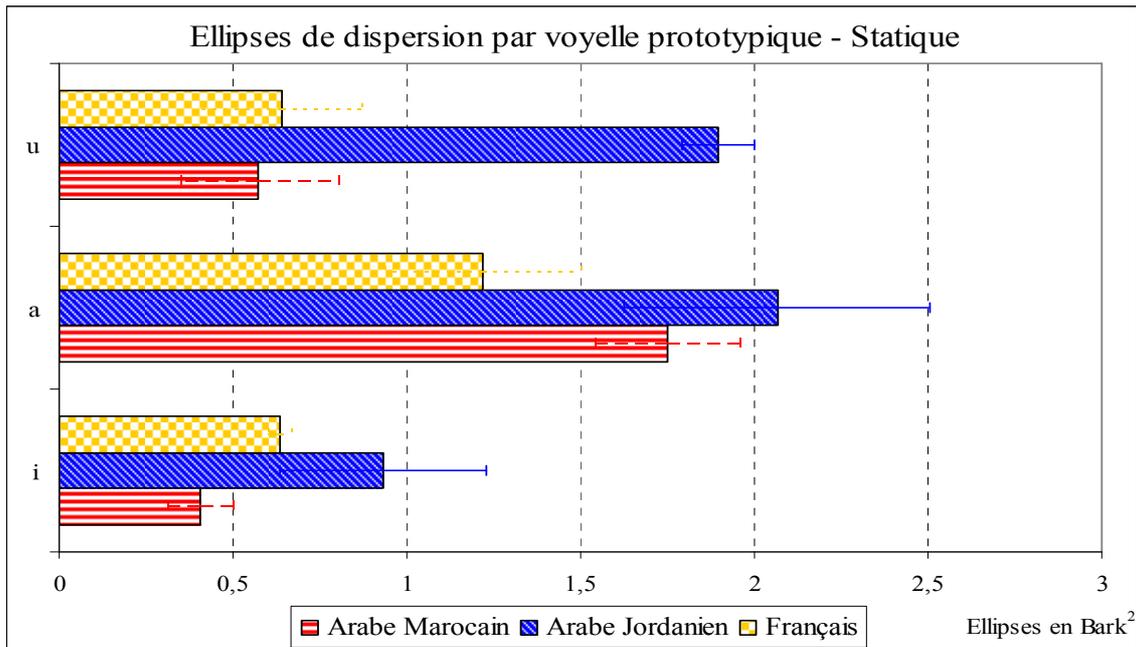


Figure 154 : Ellipses de dispersion par **prototype** dans les trois langues en **statique**.

La taille des ellipses de chaque catégorie vocalique dépend du contexte de présentation (voir Figure 155). En effet, en contexte bilabial, les ellipses de dispersion des catégories vocaliques sont plus grandes en AJ, intermédiaires en AM et petites en FR (sauf pour /i/). En contexte dental et vélaire, nous obtenons des ellipses de dispersion plus grande en AJ, intermédiaires en AM et petites en FR (sauf pour /a/).

Nous observons également que la voyelle /i/ obtient les ellipses de dispersion les moins importantes, /a/ les plus grandes et /u/ des aires intermédiaires. Ceci va dans le sens des propositions de la **théorie quantique** où les voyelles /i u/, du fait qu'elles soient quantiques, présentent les ellipses de dispersion les plus petites ; la voyelle /a/ quant à elle est variable car dépendante du degré d'aperture de la mâchoire (et donc perceptivement, les auditeurs utilisent potentiellement les connaissances qu'ils ont de leur système de production pour « retrouver » les positions favorisées pour chaque voyelle).

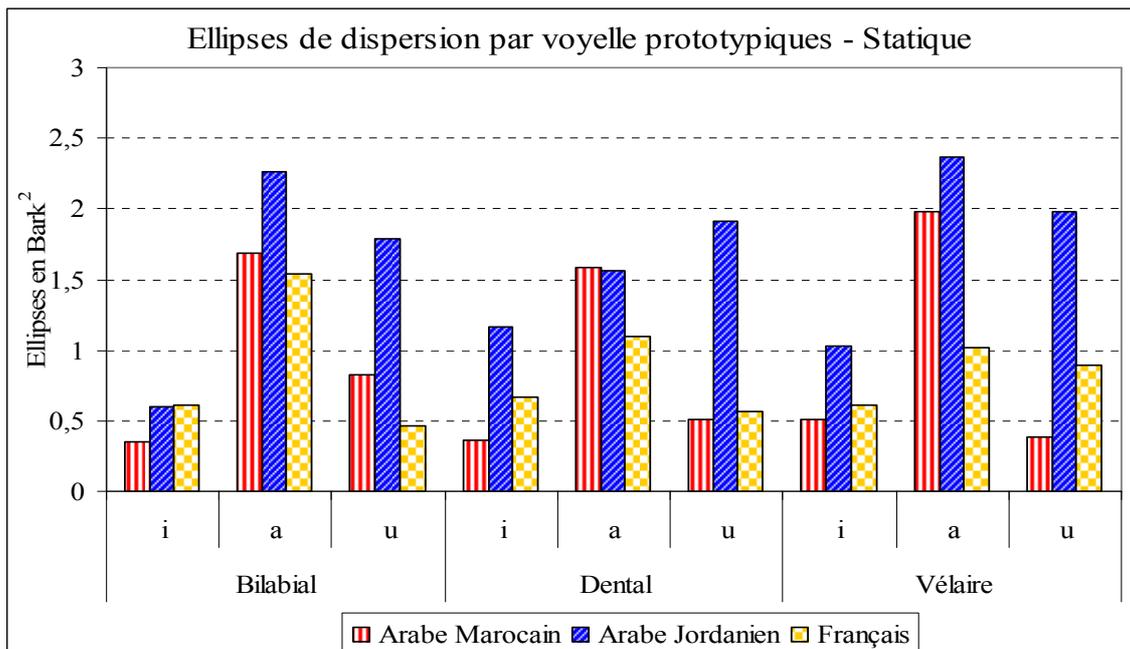


Figure 155 : Ellipses de dispersion par **prototype** dans les trois langues, en **statique**, en fonction du contexte de présentation.

Ces résultats montrent que la densité des systèmes vocaliques affecte la dispersion des voyelles **prototypiques** /i a u/ en perception de la parole. La comparaison entre les trois langues montre que c'est l'AM qui présente les aires de dispersion les plus petites comparé à l'AJ ou au FR (bien que la différence entre l'AM et l'AJ ne soit pas significative). En effectuant le calcul de taille de l'aire en fonction de **toutes** les voyelles **prototypiques** identifiées dans les trois langues, nous n'observons pas de différences significatives entre l'AM et l'AJ ; la comparaison avec le FR s'est révélée significative ($p < 0,001$) ; la tendance, elle, est inversée : c'est l'AJ qui présente la taille de l'aire la plus petite et le FR, la plus grande (avec une moyenne de 16,23 Bark² en AJ ; 17,27 Bark² en AM et 26,54 Bark² en FR). Les aires obtenues en utilisant **toutes** les voyelles **prototypiques** de chaque système sont significativement plus grandes que celles obtenues en utilisant seulement les trois voyelles communes /i a u/ (voyelles longues en arabe). Les **prototypes** identifiés par les auditeurs français sont significativement plus *Hyper-Articulés* (Lindblom, 1990 ; Johnson *et al.*, 1993, 2004) que ceux obtenus en AM ou en AJ ; aucune différence liée à la dispersion maximale des espaces vocaliques de l'AM et de l'AJ n'est vérifiée. Les différences de taille des espaces vocaliques et des différences de position des voyelles **prototypiques** /i a u/ identifiées par les auditeurs des trois langues, ne sont confirmées qu'en FR, allant ainsi à l'encontre de

l'hypothèse de stabilité de l'espace vocalique perceptif entre les langues (Meunier, 2005). Ces résultats vont dans le sens des propositions de la **théorie de la dispersion adaptée** (Lindblom, 1986), car la taille des espaces vocaliques obtenue dans les trois langues semble dépendre de la densité des systèmes. Les différences non significatives obtenues lors de la comparaison entre les espaces vocaliques en AM et en AJ ne vont pas dans le sens de l'hypothèse d'une influence de la densité des systèmes.

5.3.5.2. Résultats en dynamique (CV Target)

Les résultats graphiques de la comparaison entre les espaces vocaliques perceptifs obtenus à partir des trois voyelles **prototypiques** /i a u/ dans les trois langues sont présentés sur la Figure 156. Les espaces vocaliques perceptifs sont significativement différents dans les trois langues : l'espace vocalique en FR est significativement plus « périphérique » ; en AM, il est plus « centralisé », sur les deux axes (sur F_1 : $F(2, 1323)=11,58$; $p<0,001$; sur F_2 : $F(2, 1323)=80,85$; $p<0,001$). Nous observons des positions significativement différentes entre les trois voyelles : sur F_1 , la voyelle /i/ est significativement plus « fermée » en FR qu'en AM ou en AJ ; la voyelle /u/ est significativement plus « fermée » en FR et plus « ouverte » en AM et la voyelle /a/ est significativement plus « ouverte » en FR et plus « fermée » en AM qu'en AJ ($F(4, 1323)=151,00$; $p<0,001$) ; sur F_2 , la voyelle /i/ est significativement plus « antérieure » en FR et en AJ qu'en AM ; /a/ est significativement plus « antérieure » en AM et plus « postérieure » en FR qu'en AJ ; et /u/ est significativement plus postérieure en FR et en AJ qu'en AM ($F(4, 1323)=103,50$; $p<0,001$).

L'interaction entre les langues et les contextes de présentation n'est pas significative (voir Figure 157). Ceci indique que les espaces vocaliques obtenus dans les trois contextes de présentation sont significativement non différents en ce qui concerne les effets de « périphérisation » de l'espace en FR et la « centralisation » en AM.

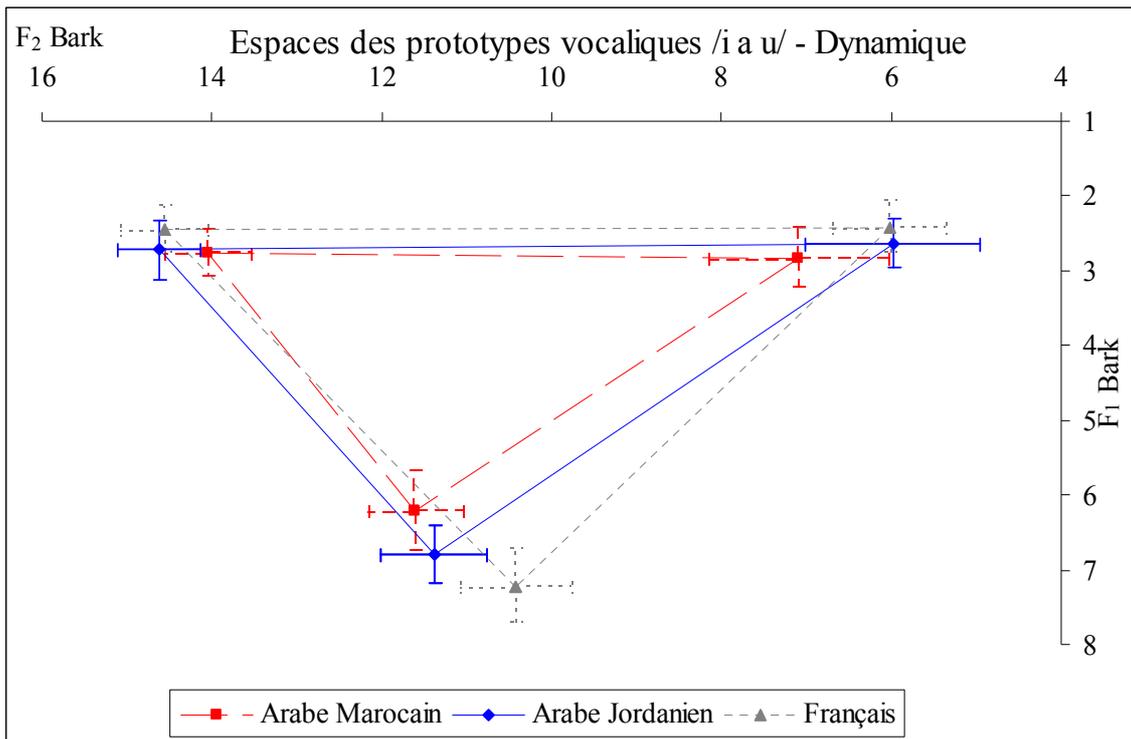


Figure 156 : Espaces des **prototypes** vocaliques /i a u/ dans les trois langues en **dynamique**.

Les résultats statistiques montrent que l'interaction entre les langues, les contextes de présentation et les voyelles est significative. L'examen attentif des graphiques présentés sur la Figure 157 montre que pour les trois contextes de production et sur F_1 , nous obtenons une tendance à ce que la voyelle /i/ soit plus « fermée » en FR qu'en AM ou en AJ (sauf en contexte /k/ où elle est plus ouverte en AJ) ; à ce que la voyelle /a/ soit plus « ouverte » en FR qu'en AM ou en AJ ; et à ce que la voyelle /u/ soit plus « fermée » en FR qu'en AM ou en AJ ($F(8, 1323)=1,71$; $p=0,092$). Sur F_2 , et pour les trois contextes de présentation, nous obtenons une tendance à ce que la voyelle /i/ soit plus « antérieure » en FR et en AJ qu'en AM ; à ce que la voyelle /a/ soit plus « antérieure » en AM et en AJ qu'en FR ; et à ce que la voyelle /u/ soit plus « antérieure » en AM qu'en AJ ou en FR (sauf en contexte /d/ où elle est plus « postérieure » en FR), ($F(8, 1080)=1,72$; $p=0,088$).

RÉSULTATS EN PERCEPTION DES VOYELLES

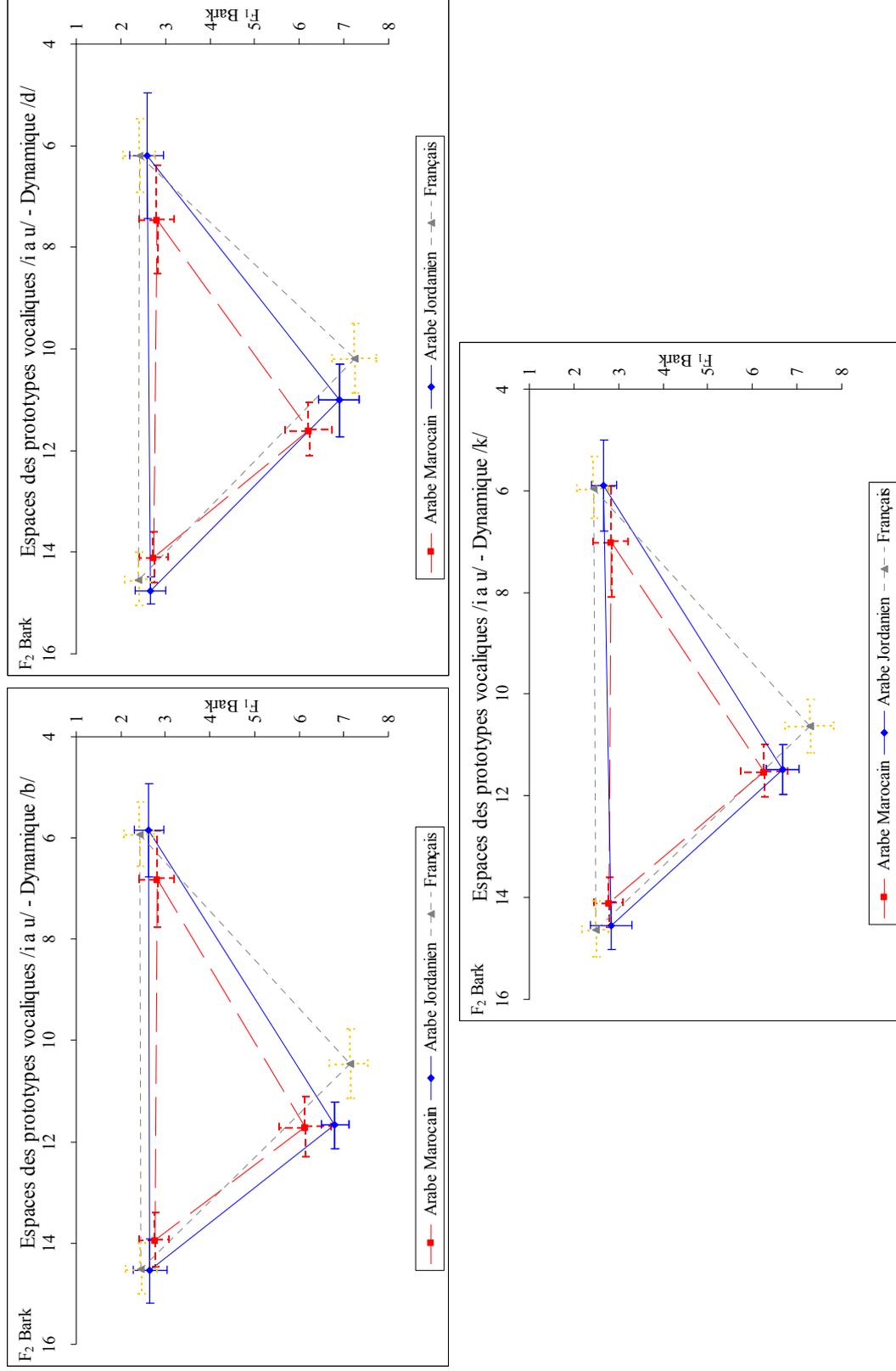


Figure 157 : Espaces des **prototypes** vocaliques /i a u/ dans les trois langues en **dynamique**, par contexte de présentation.

L'aire de dispersion moyenne en AM est significativement plus petite que celle en FR, et l'AJ obtient une dispersion intermédiaire (11,92 Bark² en AM, 17,83 Bark² en AJ et 20,50 Bark² en FR : $p < 0,001$). Les aires de dispersion en fonction du contexte de présentation montrent la même tendance (voir Figure 158). Nous obtenons sensiblement la même dispersion quel que soit le contexte de présentation. Ces résultats montrent que l'utilisation des traits **dynamiques** dans l'identification des **prototypes** vocaliques des voyelles /i a u/ a affecté la position de celles-ci dans les trois langues et quel que soit le contexte de présentation (voir Figure 157), car à notre sens, les stimuli **dynamiques** ont permis aux auditeurs des trois langues de retrouver les voyelles *Hyper-Articulées* correspondant aux **prototypes** de leurs systèmes. Cet effet est également visible pour ce qui est de la taille des espaces vocaliques. En effet, lorsque les traits **dynamiques** sont accessibles aux auditeurs, ceux-ci montrent des motifs d'identification des **prototypes** compatibles avec l'hypothèse d'une influence des tailles des espaces vocaliques (AM < AJ < FR).

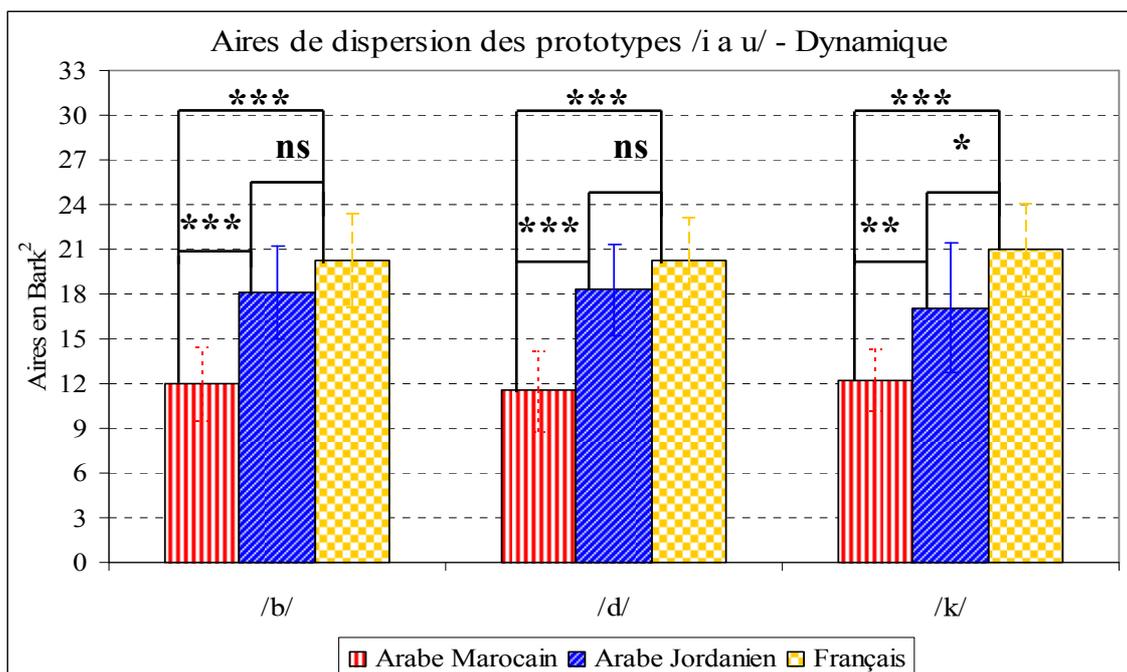


Figure 158 : Aires de dispersion des voyelles **prototypiques** /i a u/ dans les trois langues en **dynamique**, en fonction du contexte de présentation. (ns = non significative ; * = $p < 0,05$; ** = $p < 0,005$; *** = $p < 0,001$).

Comme en **statique**, nous avons examiné les effets de la densité des systèmes vocaliques sur la taille des ellipses de chacun des **prototypes** /i a u/. Les résultats graphiques

pour les trois **prototypes** sont présentés sur la Figure 159, et ceux pour les trois **prototypes** en fonction du contexte de présentation sur la Figure 160.

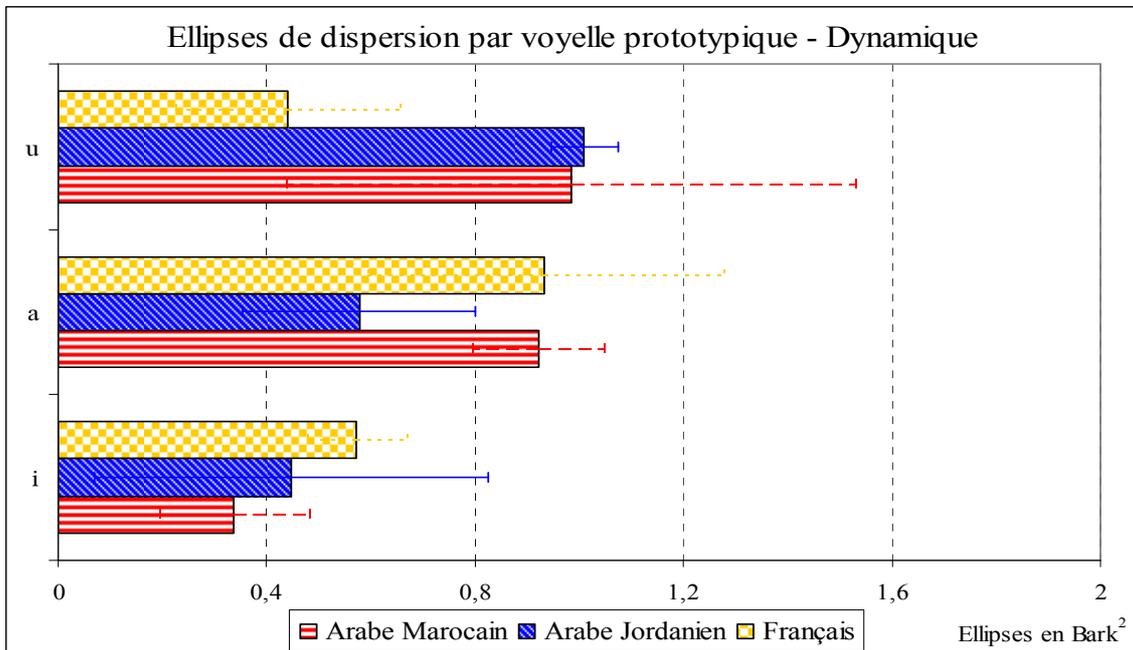


Figure 159 : Ellipses de dispersion par **prototype** dans les trois langues en **dynamique**.

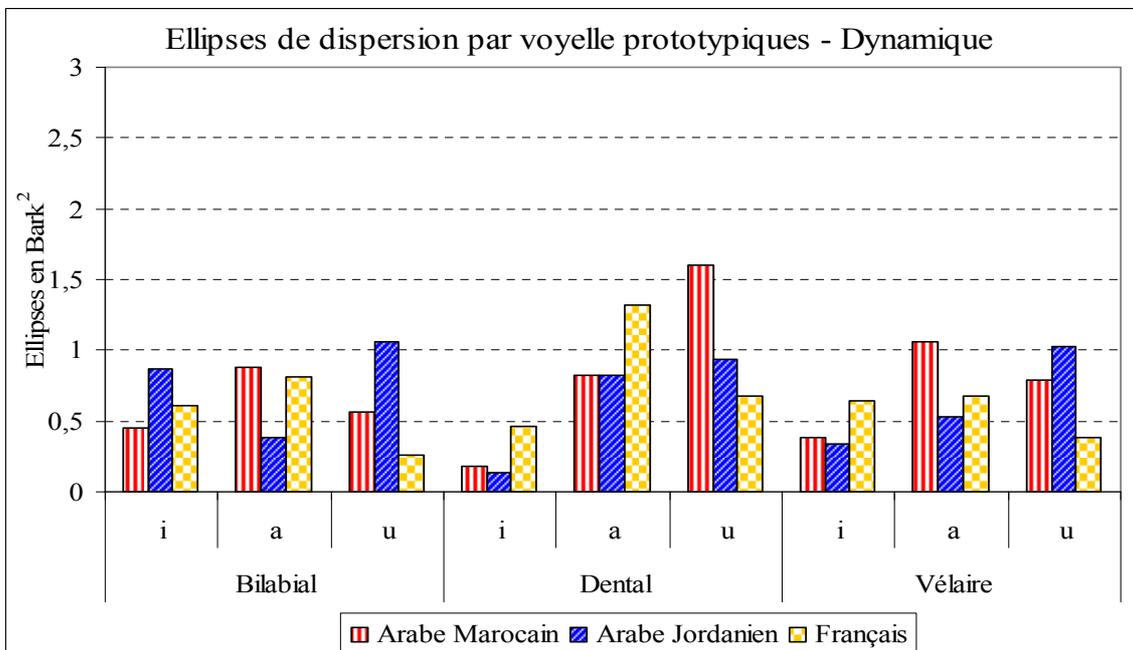


Figure 160 : Ellipses de dispersion par **prototype** dans les trois langues, en **dynamique**, en fonction du contexte de présentation.

La voyelle /i/ obtient l'ellipse moyenne la plus petite en AM et la plus grande en FR ; /a/ obtient des ellipses plus petites en AJ et plus grandes dans les deux autres langues ; et /u/

obtient des ellipses plus grandes en AM et AJ et plus petites en FR. Pour les ellipses de dispersion des **prototypes** en fonction du contexte de présentation, les résultats montrent qu'en AM et en FR, le contexte dental est celui qui présente le plus de variabilité (à l'exception de /a u/), tandis qu'en AJ, la variabilité maximale est atteinte pour /u/ quel que soit le contexte.

Les résultats obtenus en **dynamique** sont homogènes avec ceux obtenus en **statique** : ils sont même davantage conformes avec l'hypothèse d'une influence des tailles des espaces vocaliques. En revanche, les ellipses ne suivent pas l'hypothèse d'une variabilité décroissante en fonction de la taille des espaces. Nos résultats montrent une dispersion moins importante en AM, intermédiaire en AJ et plus importante en FR. Dans la partie qui suit, nous effectuerons une comparaison entre les voyelles **prototypiques** /i a u/ dans les trois langues en fonction des traits utilisés. Nous vérifierons ainsi l'apport de la **dynamique** par rapport aux traits **statiques** dans l'identification des **prototypes** dans les trois langues.

5.3.5.3. Comparaison des résultats en statique vs en dynamique

La représentation graphique de la comparaison **statique** vs **dynamique** en perception des voyelles **prototypiques** /i a u/ dans les trois langues est présentée sur la Figure 161. L'examen des espaces vocaliques présentés sur cette figure montre que ceux-ci sont plus « centralisés » en AM par rapport aux deux autres langues en fonction des deux modalités et ce sur les deux axes ; le FR présente un espace vocalique plus « périphérique » qu'en AM ou en AJ (sur F_1 , $F(2, 2644)=33,26$; $p<0,001$ et sur F_2 , $F(2, 2644)=107,95$; $p<0,001$). Les différences selon les voyelles se révèlent significatives sur les deux axes : les espaces vocaliques perceptifs obtenus en **dynamique** sont significativement plus grands que ceux obtenus en **statique** sur les deux axes (sur F_1 : $F(2, 2644)=12,74$; $p<0,001$) ; et sur F_2 : $F(2, 2644)=60,74$; $p<0,001$). L'interaction entre les langues, les traits utilisés et les voyelles est significativement différente sur les deux axes. Ainsi, sur F_1 , les espaces vocaliques obtenus en **dynamique** sont significativement plus « fermés » pour les voyelles /i u/ et plus « ouverts » pour /a/ ($F(4, 2644)=7,50$; $p<0,001$) ; sur F_2 : les espaces vocaliques perceptifs obtenus en

dynamique sont significativement plus « antérieurs » pour les voyelles /i a/ et plus « postérieurs » pour /u/ (sauf pour l'AM où c'est en **statique** que la voyelle /u/ est « postériorisée »), ($F(4, 2644)=13,15$; $p<0,001$). En ce qui concerne la taille des aires de dispersion des espaces vocaliques obtenus pour les **prototypes** /i a u/ dans les trois langues, les graphiques présentés sur la Figure 162 montrent que c'est le FR qui obtient les aires les plus grandes dans les deux modalités ($p<0,001$) ; l'AJ et l'AM obtiennent des aires non significativement différentes en **statique** ($F(1, 58) = 1,24$; $p=0,24$), en revanche, en **dynamique**, l'AJ obtient une aire de dispersion plus grande que celle en AM ($F(1, 58) = 59,29$; $p<0,001$). Les résultats obtenus dans les deux modalités en FR vont à l'encontre de la stabilité de l'espace vocalique perceptif, car la taille de l'aire obtenue est significativement plus grande que celle obtenue pour les deux langues. En ce qui concerne la comparaison entre l'AM et l'AJ, nous avons montré qu'en **statique**, les aires de dispersion sont significativement non différentes, conformément aux propositions de Meunier (2005) pour ce qui est de la stabilité de l'espace. En **dynamique**, en revanche, nous obtenons des différences significatives entre les espaces vocaliques en AM et AJ, allant ainsi à l'encontre de l'hypothèse de stabilité des espaces vocaliques perceptifs. Ces résultats vont dans le sens de l'hypothèse d'un effet de la densité des systèmes vocaliques sur la dispersion et la taille des espaces vocaliques et plus particulièrement en **dynamique**, car l'effet est vérifié dans les trois langues ; en **statique** en revanche, cette hypothèse n'est confirmée que pour le FR.

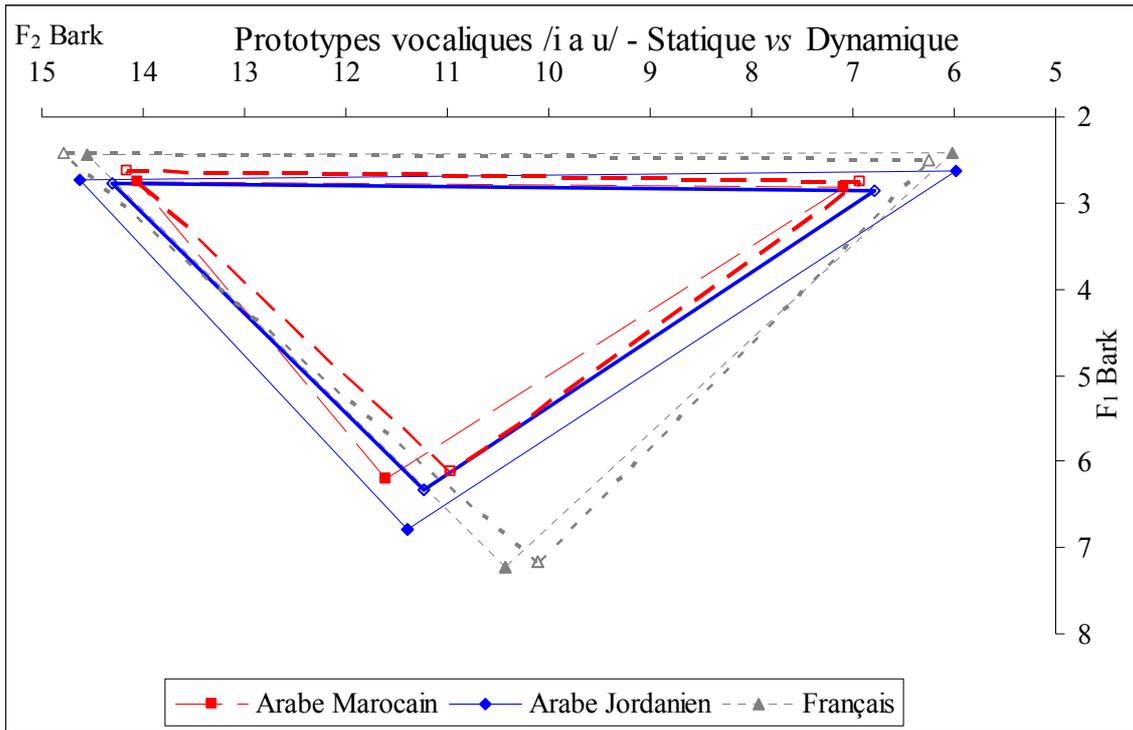


Figure 161 : **Prototypes** vocaliques de /i a u/, en **statique** (fond vide) vs en **dynamique** (fond plein), dans les trois langues.

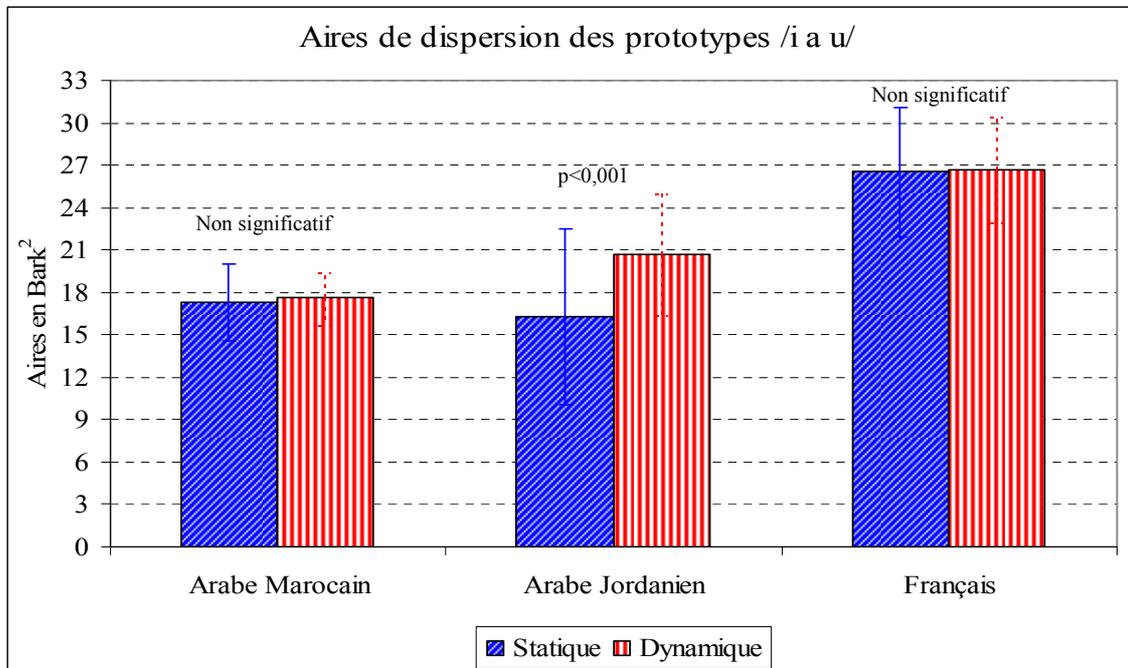


Figure 162 : Aires de dispersion des **prototypes** /i a u/ obtenues en **statique** vs **dynamique** dans les trois langues.

Afin de vérifier la validité de ces résultats quels que soient les contextes de présentation et les traits utilisés, nous présentons sur la Figure 163 les espaces vocaliques des **prototypes** /i a u/ obtenus dans les trois langues et sur la Figure 164 la taille des aires de dispersion de ces trois voyelles **prototypiques** en fonction des traits utilisés et des contextes de présentation.

En ce qui concerne la dispersion des voyelles moyennes présentée sur la Figure 163, les résultats statistiques montrent que l'interaction entre les langues, les traits utilisés et les contextes de présentation n'est pas significative, car le même effet est toujours obtenu : un espace vocalique est significativement plus « grand » lorsque la **dynamique** est utilisée dans les trois langues, quel que soit le contexte de présentation.

Pour ce qui est des résultats des aires de dispersion des voyelles **prototypiques** /i a u/ en fonction des traits utilisés et des contextes de présentation, présentées sur la Figure 164, les analyses statistiques montrent que c'est en FR que les aires sont les plus grandes quels que soient la modalité (**statique** ou **dynamique**) et le contexte de présentation ; l'AM présente les aires les plus petites (voir également les résultats obtenus pour chaque trait dans la comparaison inter-langue), ($F(2, 162)=73,70$; $p<0,001$). La comparaison des aires de dispersion des voyelles en fonction de chaque langue montre qu'en AJ, toutes les aires de dispersion sont plus grandes en **dynamique** qu'en **statique** et ce à travers les trois contextes de présentation. En AM à l'inverse, seule l'aire en contexte dental est plus grande en **statique** qu'en **dynamique**, alors qu'en contextes bilabial et vélaire, aucune différence n'est constatée. En FR, nous obtenons des aires de dispersion plus grandes en contexte bilabial et vélaire en **dynamique** par rapport au **statique**.

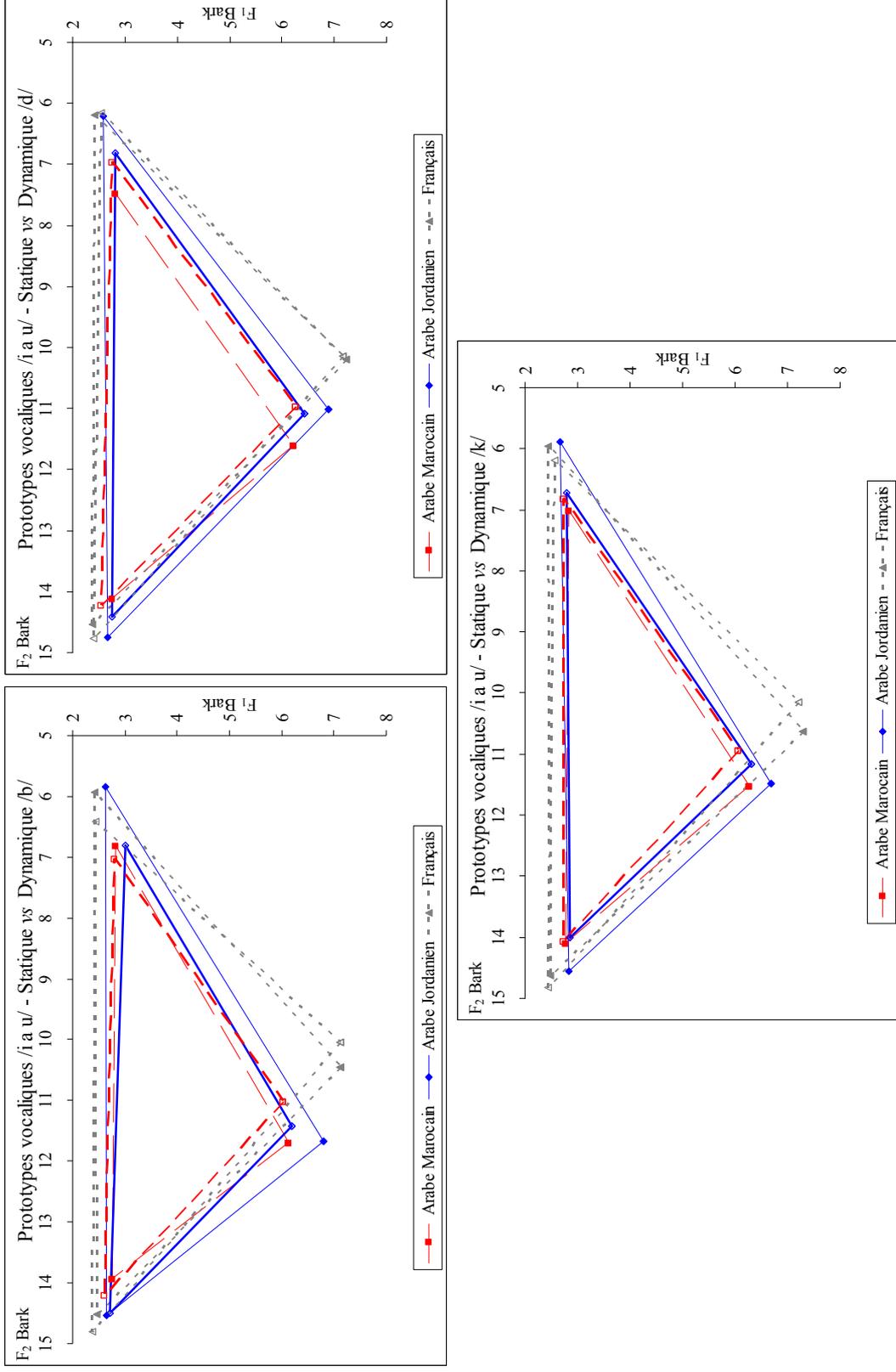


Figure 163 : Prototypes vocales de /i a u/ obtenus en **statique** (fond vide) vs **dynamique** (fond plein), dans les trois langues, en fonction du contexte de présentation.

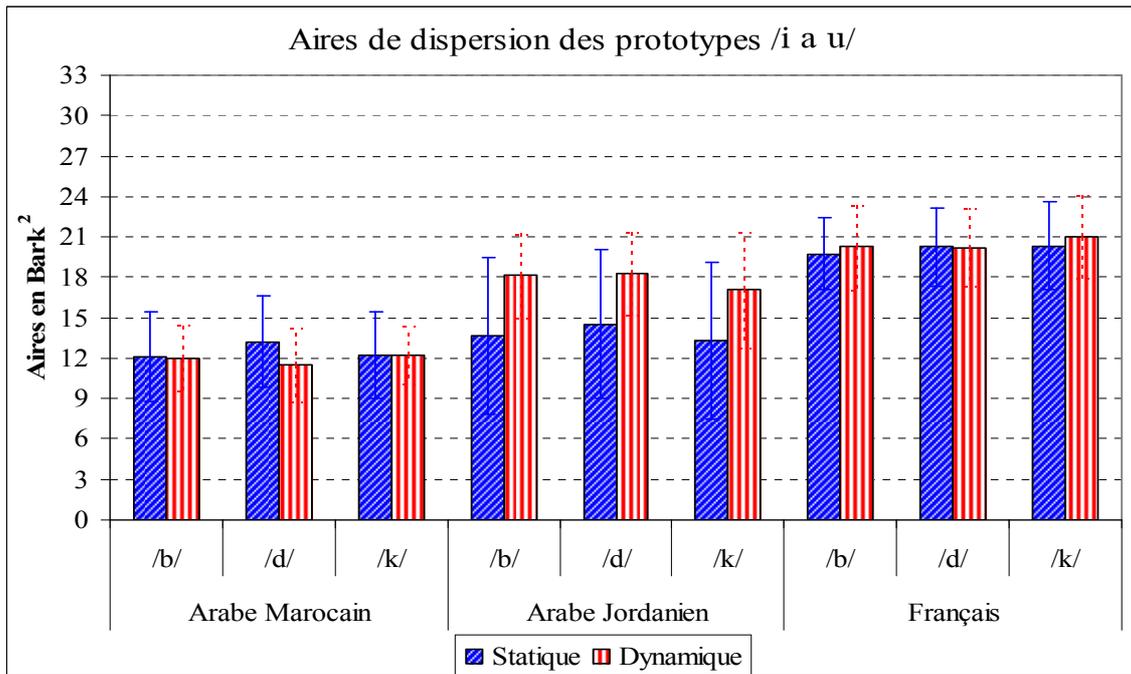


Figure 164 : Aires de dispersion des voyelles **prototypiques** /i a u/ identifiées en **statique** vs **dynamique** dans les trois langues, en fonction du contexte de présentation.

Afin de vérifier l'apport de la **dynamique** sur la dispersion de chaque voyelle **prototypique**, nous présentons sur la Figure 165 la taille des ellipses de dispersion moyennes de chaque voyelle. Les résultats montrent qu'en moyenne, les ellipses de dispersion intra-catégorie vocalique les plus importantes sont obtenues en **statique** plutôt qu'en **dynamique**. Cette tendance est vérifiée pour toutes les langues et tous les contextes de présentation, à l'exception de la voyelle /a/ en AM. Nous obtenons également des ellipses, en moyenne, plus grandes en AJ, intermédiaires en AM et plus petites en FR.

Ces tendances restent identiques lorsqu'on compare la taille des ellipses pour chaque voyelle en fonction des traits **statique** vs **dynamique** et du contexte de présentation (voir Tableau 49). Nous observons des ellipses plus grandes en moyenne en **statique** qu'en **dynamique**, ainsi que des ellipses plus grandes en AJ que dans les deux autres langues. Ces résultats montrent que la **dynamique** apporte plus de précision aux auditeurs à la tâche d'identification des **prototypes** vocaliques, comme rapporté par les auditeurs eux-mêmes.

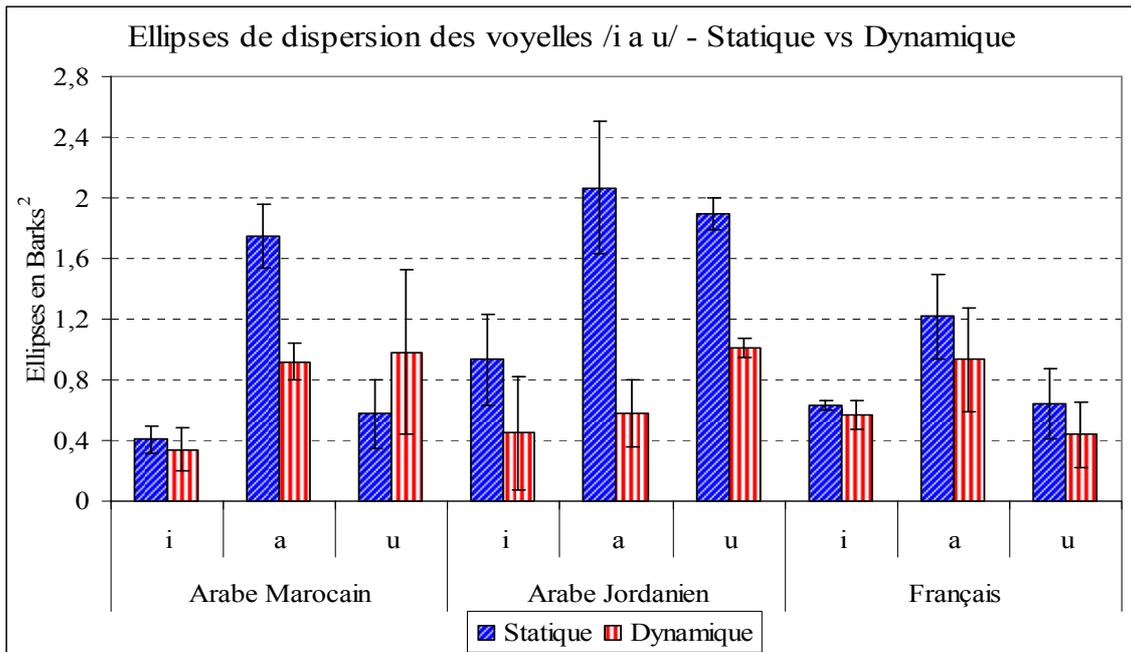


Figure 165 : Ellipses de dispersion pour chaque **prototype** /i a u/ en **statique** vs **dynamique**, dans les trois langues.

Tableau 49 : Ellipses de dispersion pour chaque **prototype** /i a u/ en **statique** vs **dynamique** dans les trois langues, en fonction du contexte de présentation (en Bark²).

		Statique			Dynamique		
		/b/	/d/	/k/	/b/	/d/	/k/
Arabe Marocain	/i/	0,35	0,36	0,51	0,45	0,18	0,38
	/a/	1,68	1,58	1,98	0,88	0,82	1,06
	/u/	0,83	0,51	0,38	0,56	1,60	0,79
Arabe Jordanien	/i/	0,60	1,17	1,03	0,87	0,13	0,34
	/a/	2,26	1,56	2,37	0,38	0,82	0,53
	/u/	1,78	1,91	1,99	1,06	0,94	1,03
Français	/i/	0,61	0,67	0,61	0,61	0,46	0,64
	/a/	1,53	1,10	1,02	0,81	1,32	0,67
	/u/	0,46	0,57	0,90	0,26	0,68	0,38

Les résultats que nous avons obtenus montrent l'existence d'une différence significative liée à la densité des systèmes qui semble affecter directement la taille des aires de dispersion des espaces vocaliques obtenus pour les trois voyelles **prototypiques** /i a u/ et des ellipses de chaque voyelle. La **dynamique** semble aider davantage les auditeurs des trois langues à « mieux » identifier les trois voyelles **prototypiques** de leurs systèmes.

Comme nous l'avons souligné, il existe des différences significatives entre les trois voyelles **prototypiques** identifiées dans les trois langues en **statique** vs **dynamique**, différences validées par les résultats de l'analyse discriminante (voir *infra*). Nous avons choisi expressément de comparer les trois voyelles **prototypiques** /i a u/ identifiées par les auditeurs des trois langues, car ce sont les trois voyelles qui délimitent l'espace vocalique et ce sont les seules voyelles communes aux trois systèmes. Nos résultats ont montré la forte « centralisation » de l'espace vocalique perceptif en AM comparé à la grande « périphérisation » de celui-ci en FR et plus spécifiquement lorsque la **dynamique** est utilisée.

Nous testons dans la partie qui suit les résultats de la classification des voyelles dans les trois langues obtenus à l'aide d'une analyse discriminante ; cette analyse nous permettra d'une part, de valider les résultats que nous venons de présenter, et d'autre part, de vérifier l'apport de la **dynamique** dans la classification des trois voyelles **prototypiques** /i a u/.

5.3.5.4. Résultats de l'analyse discriminante

Les paramètres utilisés pour cette analyse sont : les valeurs formantiques de F₁ et de F₂ des voyelles **prototypiques** identifiées par les auditeurs des trois langues en fonction des contextes de présentation et traits **statiques** vs **dynamiques**. Les taux de classification des trois langues, en fonction des traits utilisés, des contextes de présentation et des voyelles /i a u/, sont présentés dans le Tableau 50. Nous observons une augmentation significative du taux moyen de classification correcte de 7% en **dynamique** (BinomialQ ; p<0,001).

Tableau 50 : Taux de classification des voyelles **prototypiques** identifiées dans les trois langues en **statique** vs **dynamique**, des contextes de présentation et des voyelles /i a u/. (Taux significatifs : Box's M ; p<0,001).

		/b/	/d/	/k/
Statique	/i/	58,0%	45,3%	50,0%
	/a/	59,3%	62,0%	56,7%
	/u/	50,5%	54,0%	48,7%
Dynamique	/i/	51,3%	57,3%	54,7%
	/a/	77,3%	72,0%	64,7%
	/u/	57,3%	57,3%	58,0%

Les taux moyens de classification correcte des trois langues en fonction des traits utilisés, des contextes de présentation et des voyelles /i a u/ sont bas, variant de 45% à 77%. Ceci est fortement corrélé aux confusions dans la classification de l'analyse discriminante, mais également à la non utilisation du paramètre de la durée lors de l'analyse (paramètre utilisé toujours en premier lieu par l'analyse discriminante en production). C'est pourquoi nous voulions savoir laquelle des trois voyelles favorise la classification des trois langues, et laquelle des trois langues est la mieux classifiée. Nous présentons ainsi sur les Figures 166 et 167 les résultats graphiques des confusions de classification des trois langues en **statique** vs **dynamique**, respectivement, en prenant en compte chacune des voyelles comme paramètre classificatoire et chaque contexte de présentation. La comparaison entre les confusions de classification à partir des stimuli **statiques** (voir Figure 166) et **dynamiques** (voir Figure 167) montre que c'est en **dynamique** que les confusions sont les moins importantes. Il semblerait que c'est en FR que les taux de classification correcte soient les plus importants car nous obtenons moins de confusion que pour les deux autres langues.

Ces résultats vont dans le sens des différences observées entre les trois langues en fonction soit de la dispersion des voyelles à l'intérieur de l'espace vocalique soit de la taille des aires de dispersion. Il semblerait donc que l'hypothèse de l'influence de la densité des systèmes affecte bien l'organisation des espaces vocaliques dans les langues étudiées.

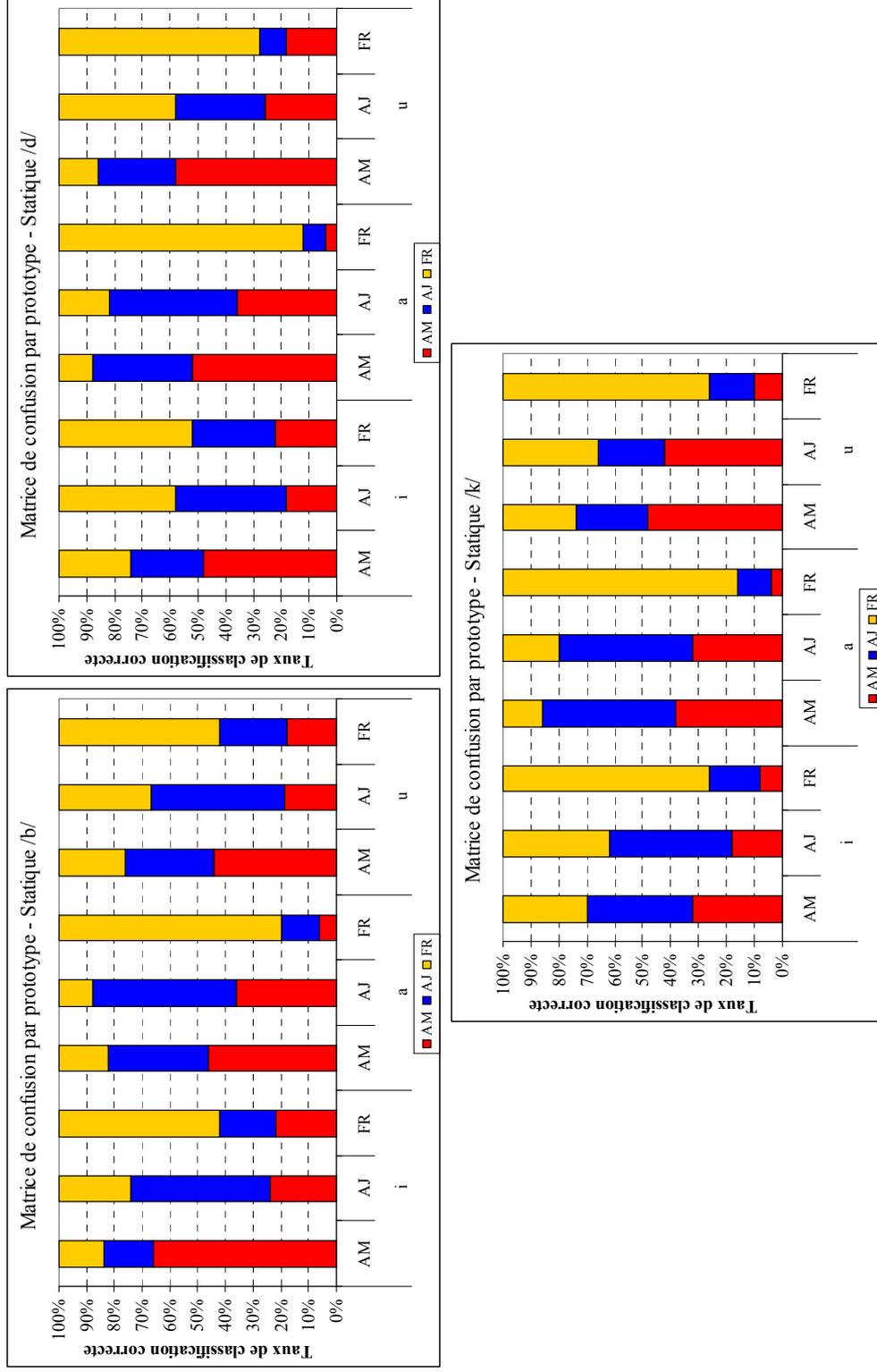


Figure 166 : Matrices de confusion de classification de chaque **prototype** /i a u/ dans les trois langues en **statique**, en fonction du contexte de présentation.

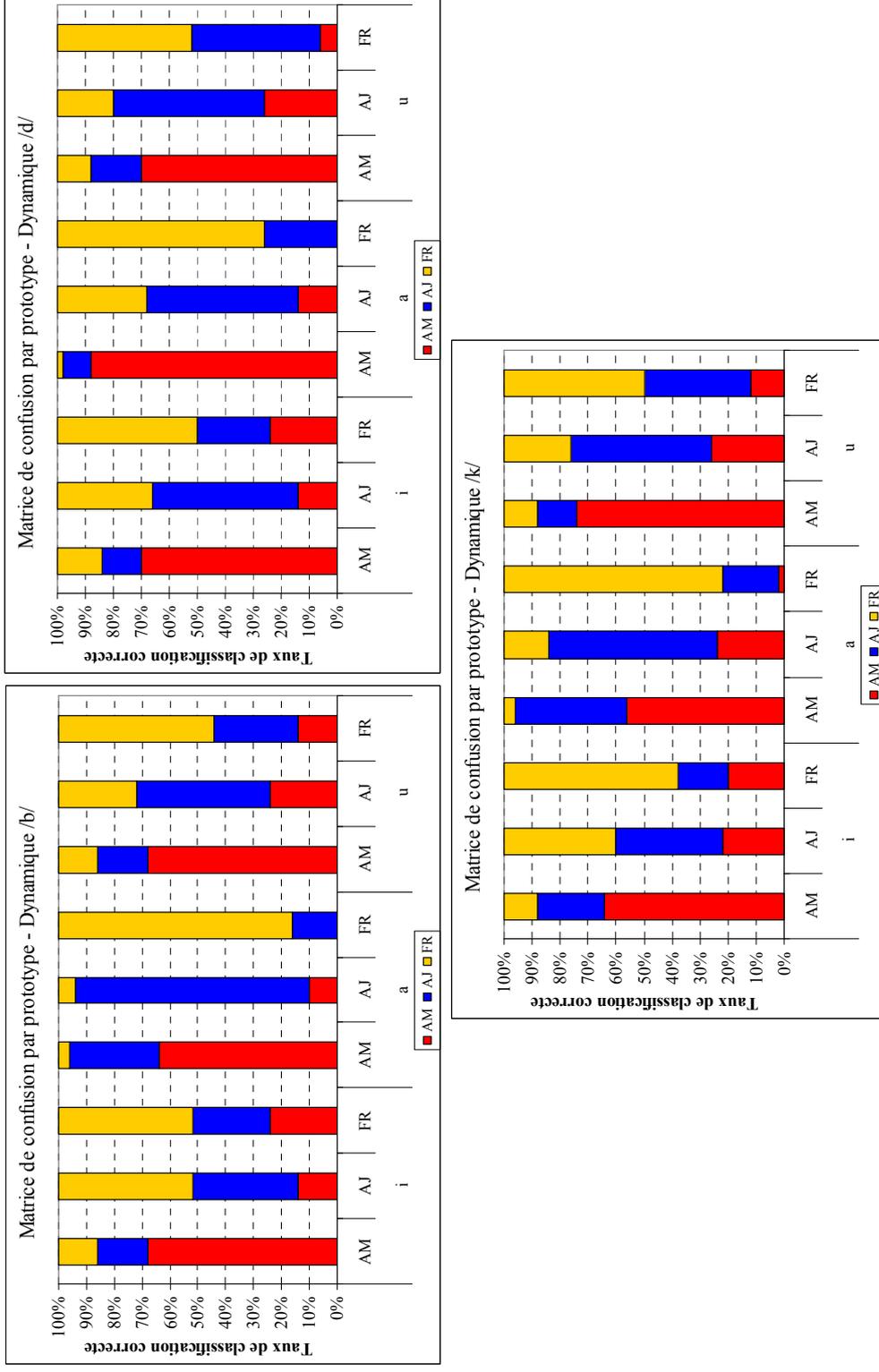


Figure 167 . Matrices de confusion de classification de chaque **prototype** /i a u/ dans les trois langues en **dynamique**, en fonction du contexte de présentation.

Nous présentons dans la partie qui suit les résultats de la seconde expérience de perception que nous avons élaborée dans cette thèse, à savoir l'expérience de Catégorisation de l'espace vocalique en utilisant les indices statiques et dynamiques.

5.4. CATÉGORISATION DE L'ESPACE VOCALIQUE

Nous cherchons dans cette expérience à savoir comment les auditeurs des trois langues catégorisent leurs espaces vocaliques. Cette expérience est une réplique de celle proposée par Hombert et Puech (1984). Nous commençons par la présentation des analyses et du traitement des données de Catégorisation, ensuite nous présentons les résultats pour chaque langue en **statique** et **dynamique**.

5.4.1. *Analyses et traitement des données*

Notre corpus de contient 28000 stimuli synthétiques (pour plus de détails sur le protocole, voir § 5.2.3.2), qui ont été classés au sein de différentes catégories vocaliques par les auditeurs des trois langues (11200 stimuli en AM, 11200 en AJ et 5600 en FR)⁷⁷. En nous basant sur les résultats obtenus par Hombert et Puech (1984), nous savons qu'une variabilité inter-individuelle importante dans la catégorisation de l'espace vocalique de chaque langue est observable. Toutefois, il est difficile de présenter les résultats individuels de cette tâche et de tirer ainsi des conclusions générales au niveau de l'organisation de chaque langue. Afin de généraliser les résultats de catégorisation que nous avons obtenus, nous avons décidé de présenter les tendances générales pour chaque langue en calculant le pourcentage des réponses de catégories pour chaque langue et des scores de réponses par système. Chaque stimulus obtient le score 1 s'il y a 3 bonnes réponses ou plus et 0 s'il y a 1 ou 2 réponses (ainsi le score maximal sera de 10 et le minimal, 0). On calcule ensuite la somme des scores. Si la somme est supérieure ou égale à 7 pour une seule catégorie, alors le stimulus est identifié, si la somme est supérieure ou égale à 7 pour plus d'une catégorie, alors le stimulus est identifié avec des confusions et si la somme est inférieure ou égale à 6, alors le stimulus

⁷⁷ 11200 stimuli en arabe = 56 stimuli * 5 répétitions * 10 auditeurs * 2 indices * 2 durées en arabe et 5600 stimuli en français = 56 stimuli * 5 répétitions * 10 auditeurs * 2 indices.

n'est pas identifié⁷⁸. Nous présentons dans cette partie les résultats de l'expérience de catégorisation des espaces vocaliques dans les trois systèmes séparément, en **statique**, puis en **dynamique**.

5.4.2. *En Arabe Marocain*

5.4.2.1. **Résultats en statiques (V Target)**

Le pourcentage des réponses des auditeurs marocains est présenté dans le Tableau 51. Nous observons une certaine facilité de catégorisation pour les stimuli représentant les trois voyelles extrêmes longues /i: a: u:/, mais une difficulté bien visible pour les trois voyelles brèves /ɪ ɐ ʊ/. Les auditeurs nous ont rapporté les difficultés qu'ils ont notées lors de la catégorisation de l'espace vocalique. Pour eux, la ressemblance entre les différents stimuli proposés peut expliquer en partie les difficultés qu'ils avaient à différencier les catégories et plus particulièrement dans le cas des trois voyelles brèves. L'examen des résultats du Tableau 51 révèle ces confusions. Afin de mieux visualiser les résultats, nous les présentons en Figure 168, en fonction des sommes des scores des bonnes réponses.

⁷⁸ Dans la présentation graphique des résultats, nous avons exclu les séries de stimuli antérieurs arrondis (i.e., stimuli N°. 3, 12, 21, 29, 37 et 43) et postérieurs non-arrondis (i.e., stimuli N°. 53, 54, 55 et 56).

Tableau 51 : Pourcentages de catégorisation des 2800 stimuli **statiques** (avec une durée de 200 ms) en AM. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	i:	ɪ	a:	ɛ	ɔ	u:	aucun	Total
1	250	2350	3100	96%						4%	100%
2	250	2150	3100	90%		2%	4%			4%	100%
3	250	1950	2900	68%	2%		10%	2%	6%	12%	100%
4	250	1750	2900	46%	2%		14%	2%	8%	28%	100%
5	250	1500	2500	16%			16%		40%	28%	100%
6	250	1250	2500	4%			18%	4%	48%	26%	100%
7	250	1050	2300				16%	2%	64%	18%	100%
8	250	850	2300		4%		8%	4%	80%	4%	100%
9	250	650	2200					6%	92%	2%	100%
10	350	2350	3100	90%		2%				8%	100%
11	350	2150	3100	88%		2%				10%	100%
12	350	1950	2900	44%	2%	6%	12%	8%	6%	22%	100%
13	350	1750	2900	22%			16%	6%	20%	36%	100%
14	350	1500	2500		2%	2%	18%	4%	58%	16%	100%
15	350	1250	2500				14%	4%	70%	12%	100%
16	350	1050	2300				12%	6%	78%	4%	100%
17	350	850	2300				4%	10%	80%	6%	100%
18	350	650	2200					6%	88%	6%	100%
19	450	2150	3100	48%	6%	28%	2%			16%	100%
20	450	1950	2900	20%	4%	30%	10%	4%	6%	26%	100%
21	450	1750	2900	16%	12%	18%	14%	4%	16%	20%	100%
22	450	1500	2500	2%	10%	8%	6%	4%	44%	26%	100%
23	450	1250	2500			6%	4%	14%	68%	8%	100%
24	450	1050	2300		2%		2%	10%	62%	24%	100%
25	450	850	2300		2%			4%	68%	26%	100%
26	450	650	2200					8%	60%	32%	100%
27	550	2150	3100	28%	4%	52%				16%	100%
28	550	1950	2900	26%	6%	54%		4%		10%	100%

N° Stimuli	F ₁	F ₂	F ₃	i:	ɪ	a:	ɛ	ɔ	u:	aucun	Total
29	550	1750	2900	6%	12%	52%	4%	6%	6%	14%	100%
30	550	1500	2500		14%	48%	8%	8%	16%	6%	100%
31	550	1250	2500		6%	46%	8%	4%	26%	10%	100%
32	550	1050	2300		8%	50%		4%	22%	16%	100%
33	550	850	2300		4%	36%		6%	30%	24%	100%
34	550	650	2200		2%	42%		2%	22%	32%	100%
35	650	1950	2900	6%	4%	62%	8%	2%		18%	100%
36	650	1750	2900		16%	78%	2%			4%	100%
37	650	1500	2500		8%	76%		8%		8%	100%
38	650	1250	2500		4%	84%		2%		10%	100%
39	650	1050	2300		6%	84%			2%	8%	100%
40	650	850	2300		6%	70%			4%	20%	100%
41	750	1950	2900		8%	74%	2%			16%	100%
42	750	1750	2900		8%	80%	2%			10%	100%
43	750	1500	2500		4%	84%				12%	100%
44	750	1250	2500			84%				16%	100%
45	750	1050	2300		6%	84%				10%	100%
46	750	850	2300		2%	84%				14%	100%
47	250	1950	2300	58%			10%			22%	100%
48	350	1950	2300	60%	4%		8%	2%	2%	24%	100%
49	450	1750	2300	12%	10%	18%	16%	4%	18%	22%	100%
50	550	1750	2300	2%	18%	58%	6%	8%		8%	100%
51	650	1500	2300		10%	74%	4%	4%		8%	100%
52	750	1500	2300		6%	78%	4%	2%		10%	100%
53	250	850	2700		2%		6%		88%	4%	100%
54	350	850	2700					6%	90%	4%	100%
55	450	850	2700		4%	4%		2%	60%	30%	100%
56	550	850	2700		6%	42%			26%	26%	100%

Nous observons que les trois voyelles longues /i: a: u:/ ont été catégorisées majoritairement par les auditeurs marocains. La voyelle /i:/ a été catégorisée par les auditeurs marocains à partir de quatre stimuli différents concentrés dans l'extrémité de l'espace vocalique (N°. 1, 2 10 et 11) avec plus de 85% de chance ; /a:/ est catégorisée par douze stimuli différents couvrant la totalité de l'axe avant~arrière et /u:/ est catégorisée à partir de neuf stimuli différents sur l'axe postérieure. Les résultats de catégorisation pour la voyelle /u:/ sont conformes à la littérature et à nos résultats en production et perception : effet d'antériorisation de la voyelle /u:/ en contexte dental. Pour la catégorisation des trois voyelles brèves /ɪ ɐ ʊ/, représentant différentes voyelles /ə/, nous observons qu'aucune voyelle n'a été catégorisée majoritairement par les auditeurs marocains. La plupart de temps, les auditeurs ne savaient que choisir comme catégorie représentant ces voyelles. Nous observons également une délimitation de l'espace vocalique par la catégorie de réponse : « confondu ».

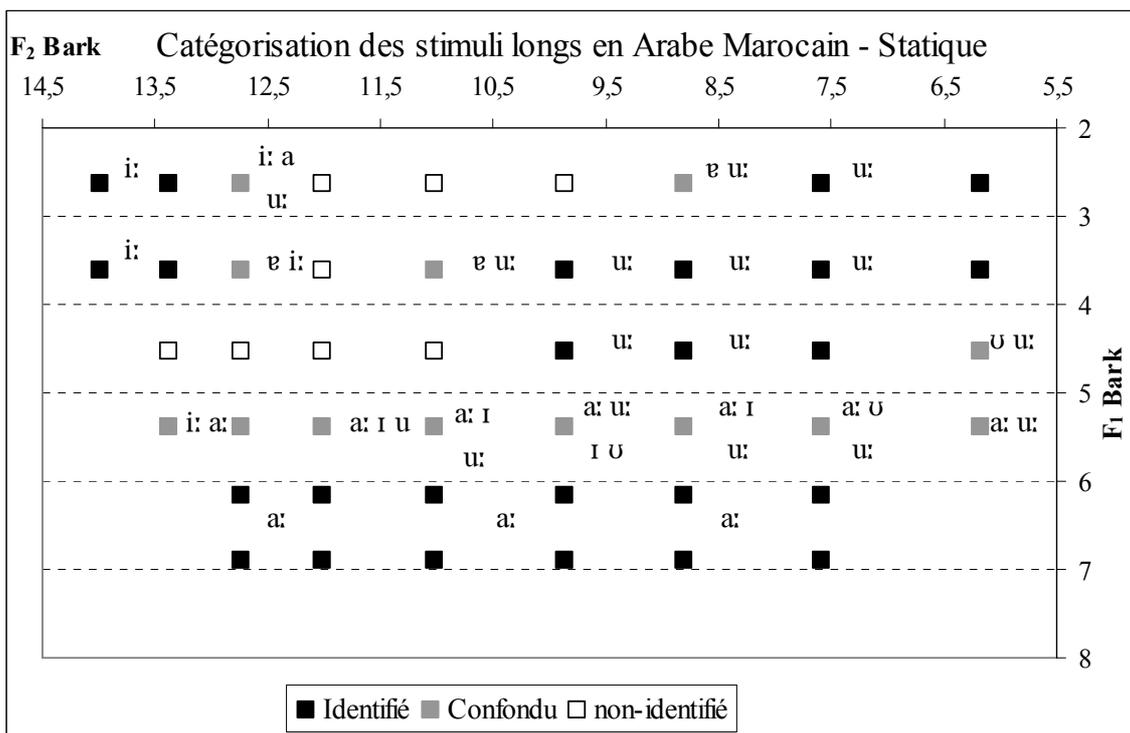


Figure 168 : Catégorisation de l'espace vocalique en AM en **statique** avec une durée de 200 ms. « Identifié » même catégorie, ≥ 7 , « Confondu » plusieurs catégories ≥ 7 , « non-identifié » = une ou plusieurs catégories ≤ 6 .

Intéressons-nous maintenant aux réponses de catégorisation des auditeurs marocains à partir des stimuli brefs (i.e., à 100 ms), (voir Tableau 52).

Tableau 52 : Pourcentages de catégorisation des 2800 stimuli **statiques** (avec une durée de 100 ms) en AM. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	i:	i	a:	e	o	u:	aucun	Total
57	250	2350	3100	54%	14%					32%	100%
58	250	2150	3100	52%	12%		6%	2%		28%	100%
59	250	1950	2900	40%	8%		12%	6%		34%	100%
60	250	1750	2900	36%	2%		18%	10%	2%	32%	100%
61	250	1500	2500	18%	2%		26%	12%	8%	34%	100%
62	250	1250	2500		2%		28%	18%	14%	38%	100%
63	250	1050	2300	2%			30%	14%	26%	28%	100%
64	250	850	2300				26%	16%	28%	30%	100%
65	250	650	2200		2%		2%	24%	42%	30%	100%
66	350	2350	3100	60%	10%		2%			28%	100%
67	350	2150	3100	48%	18%		6%			28%	100%
68	350	1950	2900	28%	8%		14%	14%		36%	100%
69	350	1750	2900	18%	6%		24%	28%		24%	100%
70	350	1500	2500	2%			20%	26%	24%	28%	100%
71	350	1250	2500				20%	20%	32%	28%	100%
72	350	1050	2300		2%		26%	20%	30%	22%	100%
73	350	850	2300				10%	26%	32%	32%	100%
74	350	650	2200					22%	40%	38%	100%
75	450	2150	3100	22%	22%	2%	14%	4%		36%	100%
76	450	1950	2900	16%	18%		24%	16%		26%	100%
77	450	1750	2900	10%	8%		20%	34%	4%	24%	100%
78	450	1500	2500	4%	4%		18%	32%	12%	30%	100%
79	450	1250	2500		6%		14%	28%	24%	28%	100%
80	450	1050	2300				16%	24%	32%	28%	100%
81	450	850	2300				4%	28%	30%	38%	100%
82	450	650	2200				4%	28%	26%	42%	100%
83	550	2150	3100	16%	18%	4%	24%	4%		34%	100%
84	550	1950	2900	12%	16%	2%	36%	4%		30%	100%
85	550	1750	2900	10%	18%		34%	16%		22%	100%
86	550	1500	2500		18%	2%	36%	16%		28%	100%
87	550	1250	2500		16%	4%	24%	24%	2%	30%	100%
88	550	1050	2300		8%	2%	22%	26%	10%	32%	100%
89	550	850	2300		4%	4%	10%	20%	12%	50%	100%
90	550	650	2200		6%	6%	14%	16%	12%	46%	100%
91	650	1950	2900	4%	14%	4%	38%	4%		36%	100%
92	650	1750	2900		12%	10%	44%	10%		24%	100%
93	650	1500	2500		14%	10%	44%	6%		26%	100%
94	650	1250	2500		12%	14%	36%	6%		32%	100%
95	650	1050	2300		16%	22%	20%	6%	4%	32%	100%
96	650	850	2300		18%	24%	10%	16%		32%	100%
97	750	1950	2900		8%	10%	32%	8%		42%	100%
98	750	1750	2900		14%	16%	22%	8%		40%	100%
99	750	1500	2500		12%	12%	28%	10%		38%	100%
100	750	1250	2500		10%	24%	18%	6%		42%	100%
101	750	1050	2300		6%	24%	18%			52%	100%
102	750	850	2300		12%	24%	20%	4%		40%	100%
103	250	1950	2300	34%	6%		12%	8%		40%	100%
104	350	1950	2300	42%	8%		16%	6%		28%	100%
105	450	1750	2300	10%	12%		26%	16%	2%	34%	100%
106	550	1750	2300	10%	20%		38%	16%		16%	100%
107	650	1500	2300		26%	18%	34%	2%		20%	100%
108	750	1500	2300		14%	18%	32%			36%	100%
109	250	850	2700				18%	20%	26%	36%	100%
110	350	850	2700				10%	26%	34%	30%	100%
111	450	850	2700				2%	30%	26%	42%	100%
112	550	850	2700		2%	6%	10%	18%	16%	48%	100%

Nous observons que seule la voyelle /i:/ obtient des pourcentages de catégorisation supérieurs à 50% (N° 57, 58 et 66). L'opposition de quantité vocalique en arabe fait que des différences de durée vocalique peuvent être l'origine de la discrimination des voyelles entre elles et plus particulièrement dans le cas des voyelles brèves. Or, les résultats de catégorisation obtenus montrent l'effet inverse. Nous observons également des confusions importantes pour toutes les catégories vocaliques. Afin de mieux visualiser ces confusions, nous présentons les résultats de catégorisation sur la Figure 169. L'examen de cette figure révèle qu'aucun stimulus n'a été catégorisé majoritairement par les auditeurs marocains. La plupart des confusions concernaient les voyelles brèves, soit avec une des voyelles longues /i: a: u:/, soit avec la réponse « aucun de correspond ».

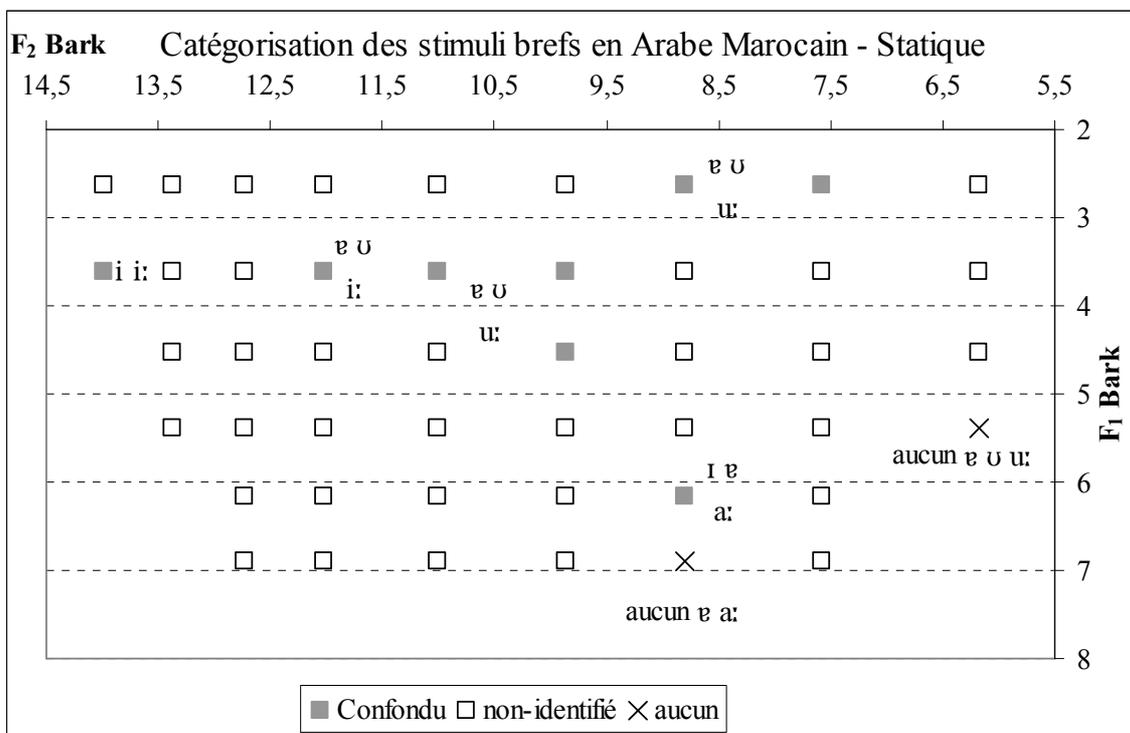


Figure 169 : Catégorisation de l'espace vocalique en AM en **statique** avec une durée de 100 ms. « Confondu » plusieurs catégories ≥ 7, « non-identifié » = une ou plusieurs catégories ≤ 6, « aucun » aucun ne correspond ≥ 7.

Nous nous intéressons dans la partie qui suit aux résultats de catégorisation de l'espace vocalique par les auditeurs marocains en **dynamique**.

5.4.2.2. Résultats en dynamique (CV Target)

Les pourcentages de catégorisation des stimuli longs en **dynamique** en AM sont présentés dans le Tableau 53. L'examen attentif des résultats montre que les auditeurs marocains sont plus précis dans leurs catégorisations. Les trois voyelles longues /i: a: u:/ sont catégorisées avec exactitude avec des pourcentages approchant les 98%. Les trois voyelles brèves /ɪ ʌ ʊ/ ne sont pas catégorisées par les auditeurs marocains. Comme avec les stimuli statiques, les auditeurs marocains nous ont révélé la difficulté qu'ils avaient à catégoriser les trois voyelles brèves dans des catégories différentes. Pour eux, en effet, ces trois voyelles sont si proches qu'ils n'arrivaient pas à les différencier. Sur la Figure 170, nous présentons les résultats de la catégorisation des voyelles en **dynamique**.

Comme avec les stimuli statiques, les trois voyelles longues /i: a: u:/ ont été majoritairement catégorisées par les auditeurs marocains. La voyelle /i:/ a été catégorisée avec quatre stimuli ; /a:/ avec quatorze stimuli et /u:/ avec sept. Les trois voyelles brèves /ɪ ʌ ʊ/ n'ont pas été identifiées par les auditeurs marocains. Par conséquent aucune voyelle brève n'a été catégorisée ni confondue.

Nous retrouvons les mêmes catégories identifiées en utilisant les deux modalités, i.e., **statique** vs **dynamique**. Il existe néanmoins des différences au niveau de la catégorisation des deux voyelles /a: u:/. En effet, la voyelle /a:/ a été catégorisée avec deux stimuli en plus qui se rapprochent du centre de l'espace vocalique conformément aux résultats de production des voyelles. Elle est confondue dans quatre autres stimuli soit avec /i:/ soit avec /u:/, marquant ainsi les frontières des catégories. La voyelle /u:/, quant à elle, est identifiée, avec moins de stimuli, comparée à sa catégorisation avec les stimuli statiques. Sa catégorisation est plus robuste car elle n'englobe pas les stimuli d'aperture médiane. Le stimulus ayant les fréquences 2,5 Bark sur F₁ et 12 Bark sur F₂ marque la frontière entre les deux voyelles /i: u:/, avec des confusions avec /ə/, représentant un schwa.

Tableau 53 : Pourcentages de catégorisation des 2800 stimuli dynamiques (avec une durée de 200 ms) en AM. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	i:	ɪ	a:	ɛ	o	u:	aucun	Total
1	250	2350	3100	98%						2%	100%
2	250	2150	3100	96%			2%			2%	100%
3	250	1950	2900	68%			10%		2%	20%	100%
4	250	1750	2900	48%			8%	2%	24%	18%	100%
5	250	1500	2500	20%	4%		10%		50%	16%	100%
6	250	1250	2500	4%			10%	2%	72%	12%	100%
7	250	1050	2300				12%	2%	74%	12%	100%
8	250	850	2300				2%		92%	6%	100%
9	250	650	2200						82%	18%	100%
10	350	2350	3100	82%		6%				12%	100%
11	350	2150	3100	80%		4%				16%	100%
12	350	1950	2900	46%	4%	2%	10%		8%	30%	100%
13	350	1750	2900	16%	4%	6%	10%	2%	36%	26%	100%
14	350	1500	2500	2%	2%		10%	4%	58%	24%	100%
15	350	1250	2500		2%		10%	2%	62%	24%	100%
16	350	1050	2300			2%	10%	6%	66%	16%	100%
17	350	850	2300					2%	82%	16%	100%
18	350	650	2200						66%	34%	100%
19	450	2150	3100	44%	2%	22%		4%		28%	100%
20	450	1950	2900	24%		38%	10%		4%	24%	100%
21	450	1750	2900	16%		12%	12%	8%	18%	34%	100%
22	450	1500	2500	2%	4%	10%	2%	6%	42%	34%	100%
23	450	1250	2500		4%	14%	4%	6%	42%	30%	100%
24	450	1050	2300		2%	6%	2%	2%	52%	36%	100%
25	450	850	2300		2%	2%			56%	40%	100%
26	450	650	2200						56%	44%	100%
27	550	2150	3100	26%		54%	2%			18%	100%
28	550	1950	2900	20%		74%				6%	100%

N° Stimuli	F ₁	F ₂	F ₃	i:	ɪ	a:	ɛ	o	u:	aucun	Total
29	550	1750	2900	2%	8%	66%	4%		6%	14%	100%
30	550	1500	2500		2%	60%			20%	18%	100%
31	550	1250	2500		2%	50%	2%		26%	20%	100%
32	550	1050	2300			42%			34%	24%	100%
33	550	850	2300			26%			34%	40%	100%
34	550	650	2200			34%			30%	36%	100%
35	650	1950	2900	12%	2%	84%				2%	100%
36	650	1750	2900	4%		92%				4%	100%
37	650	1500	2500		12%	24%	24%	12%		28%	100%
38	650	1250	2500			76%			10%	14%	100%
39	650	1050	2300			82%			6%	12%	100%
40	650	850	2300			74%		2%	10%	14%	100%
41	750	1950	2900	4%	4%	86%				6%	100%
42	750	1750	2900			92%				8%	100%
43	750	1500	2500			98%				2%	100%
44	750	1250	2500			92%				8%	100%
45	750	1050	2300			90%				10%	100%
46	750	850	2300			88%				12%	100%
47	250	1950	2300	62%			10%		6%	22%	100%
48	350	1950	2300	52%	2%		8%		6%	32%	100%
49	450	1750	2300	16%		22%	10%		6%	46%	100%
50	550	1750	2300	14%	4%	56%	4%	4%	2%	16%	100%
51	650	1500	2300		2%	92%			2%	4%	100%
52	750	1500	2300			96%	2%			2%	100%
53	250	850	2700				2%		96%	2%	100%
54	350	850	2700			2%			76%	22%	100%
55	450	850	2700			4%			50%	46%	100%
56	550	850	2700			34%			28%	38%	100%

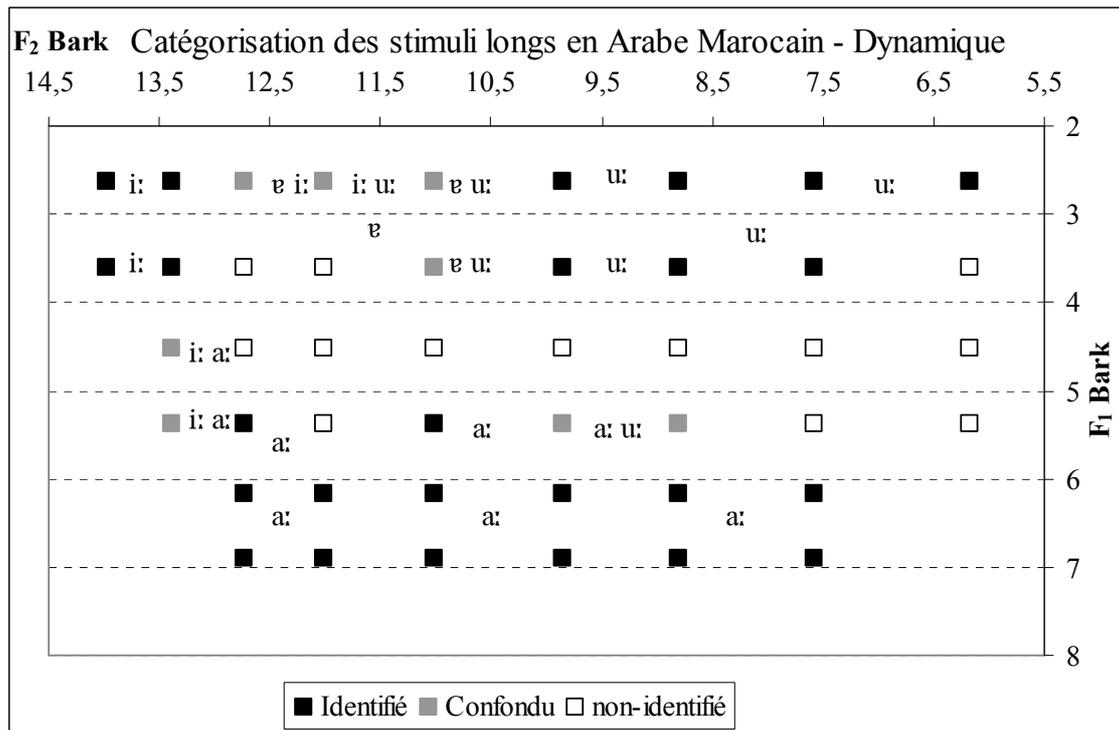


Figure 170 : Catégorisation de l'espace vocalique en AM en **dynamique** avec une durée de 200 ms. « Identifié » même catégorie, ≥ 7 , « Confondu » plusieurs catégories ≥ 7 , « non-identifié » = une ou plusieurs catégories ≤ 6 .

Intéressons-nous maintenant aux résultats de catégorisation de l'espace vocalique en AM, obtenus à partir des stimuli **dynamiques** brefs. Le Tableau 54 donne les pourcentages de catégorisation de ces stimuli. Nous pouvons observer que seule la voyelle /i:/ a été catégorisée avec plus de 50% des réponses par les auditeurs marocains à partir de trois stimuli différents. Les autres voyelles ont été confondues soit entre elles soit avec la case « aucun ne correspond ». La Figure 171 montre les résultats de la catégorisation des stimuli **dynamiques** brefs par les auditeurs marocains en fonction des scores de bonnes réponses. Nous observons que seule la voyelle /i:/ a été catégorisée correctement par au moins sept auditeurs sur dix, qui se sont basés sur les fréquences de formants et non sur la durée (résultats obtenus également pour la catégorisation en **statique**, et pour la tâche d'identification des **prototypes**, voir § 5.3.2). Les trois voyelles brèves /ɪ ʊ u/ n'ont pas été catégorisées du tout par les auditeurs marocains. Les raisons avancées par les auditeurs concernent la durée des stimuli qui, semble-t-il, n'a pas aidé à identifier les catégories vocaliques brèves. En plus, pour eux, les items utilisés se produisent de la même manière, i.e., avec une voyelle [ə], c'est pourquoi ils n'arrivaient pas à choisir entre les trois items lorsqu'ils entendaient un stimulus proche de [ə].

Tableau 54 : Pourcentages de catégorisation des 2800 stimuli dynamiques (avec une durée de 100 ms) en AM. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N°. Stimuli	F ₁	F ₂	F ₃	i:	i	a:	ɐ	o	u:	aucun	Total
57	250	2350	3100	56%	10%					34%	100%
58	250	2150	3100	56%	10%		2%			32%	100%
59	250	1950	2900	40%	4%		12%	6%		38%	100%
60	250	1750	2900	36%	4%		14%	2%	2%	42%	100%
61	250	1500	2500	8%			14%	12%	20%	46%	100%
62	250	1250	2500				22%	16%	22%	40%	100%
63	250	1050	2300				24%	24%	16%	36%	100%
64	250	850	2300				10%	14%	30%	46%	100%
65	250	650	2200				4%	22%	16%	58%	100%
66	350	2350	3100	48%	6%			2%		44%	100%
67	350	2150	3100	38%	16%		4%	2%		40%	100%
68	350	1950	2900	22%	14%		24%	8%		32%	100%
69	350	1750	2900	12%	2%		24%	22%	2%	38%	100%
70	350	1500	2500		8%		14%	18%	14%	46%	100%
71	350	1250	2500		6%		20%	20%	16%	38%	100%
72	350	1050	2300				12%	20%	14%	54%	100%
73	350	850	2300		2%		10%	22%	24%	42%	100%
74	350	650	2200				2%	20%	24%	54%	100%
75	450	2150	3100	14%	20%		14%	2%		50%	100%
76	450	1950	2900	18%	14%		20%	12%		28%	100%
77	450	1750	2900	8%	4%		26%	16%	8%	34%	100%
78	450	1500	2500		12%		24%	22%	2%	40%	100%
79	450	1250	2500		6%		8%	22%	20%	42%	100%
80	450	1050	2300		2%		4%	28%	22%	44%	100%
81	450	850	2300				4%	2%	32%	48%	100%
82	450	650	2200					22%	22%	56%	100%
83	550	2150	3100	2%	22%		14%	4%		58%	100%
84	550	1950	2900		28%		18%	4%		50%	100%

N°. Stimuli	F ₁	F ₂	F ₃	i:	i	a:	ɐ	o	u:	aucun	Total
85	550	1750	2900		20%	2%	22%	18%		38%	100%
86	550	1500	2500		8%	14%	16%	22%	8%	32%	100%
87	550	1250	2500		14%	12%	8%	24%	10%	32%	100%
88	550	1050	2300		2%	8%	10%	24%	16%	40%	100%
89	550	850	2300			8%	6%	14%	12%	60%	100%
90	550	650	2200		2%	10%	6%	10%	10%	62%	100%
91	650	1950	2900		8%	28%	24%	4%		36%	100%
92	650	1750	2900	2%	20%	26%	24%	4%		24%	100%
93	650	1500	2500		12%	22%	22%	10%		34%	100%
94	650	1250	2500		2%	34%	16%	14%		34%	100%
95	650	1050	2300		2%	32%	18%	10%	2%	36%	100%
96	650	850	2300		2%	28%	16%	8%	2%	44%	100%
97	750	1950	2900		8%	28%	24%	2%		38%	100%
98	750	1750	2900		4%	34%	32%			30%	100%
99	750	1500	2500		4%	26%	30%	4%		36%	100%
100	750	1250	2500		2%	42%	12%			44%	100%
101	750	1050	2300		4%	36%	22%			38%	100%
102	750	850	2300		6%	34%	18%			42%	100%
103	250	1950	2300	50%	2%		12%	4%		32%	100%
104	350	1950	2300	28%	8%		14%	10%		40%	100%
105	450	1750	2300	4%	14%	2%	20%	12%		48%	100%
106	550	1750	2300		14%	4%	30%	8%		44%	100%
107	650	1500	2300		12%	16%	34%	12%		26%	100%
108	750	1500	2300		6%	38%	22%			34%	100%
109	250	850	2700		2%		2%	24%	40%	32%	100%
110	350	850	2700		2%		4%	24%	30%	40%	100%
111	450	850	2700		2%			20%	32%	46%	100%
112	550	850	2700			6%		18%	14%	62%	100%

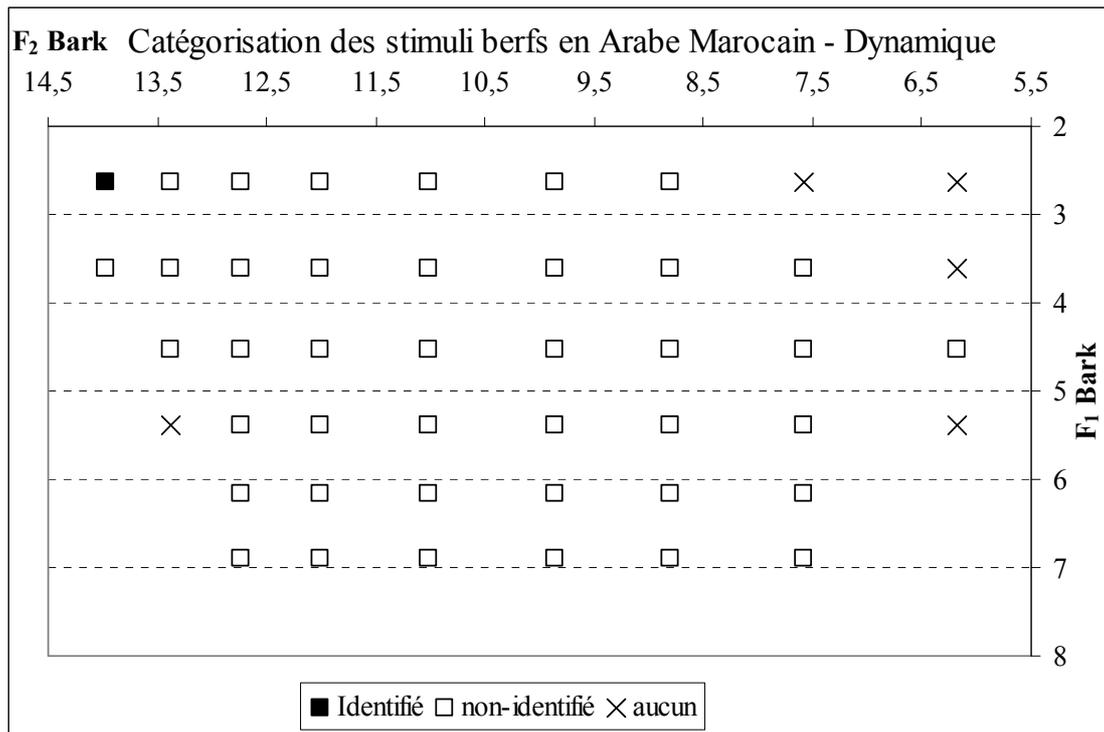


Figure 171 : Catégorisation de l'espace vocalique en AM en **dynamique** avec une durée de 100 ms. « Confondu » plusieurs catégories ≥ 7 , « non-identifié » = une ou plusieurs catégories ≤ 6 , « aucun » aucun ne correspond ≥ 7 .

Les auditeurs marocains sont capables de catégoriser les stimuli en **statique** et/ou **dynamique** en fonction des voyelles de leur(s) système(s). L'opposition de durée des stimuli proposée dans cette expérience n'a pas, semble-t-il, aidé les auditeurs marocains à différencier les voyelles longues des brèves, c'est plutôt les différences de timbres qui ont permis de différencier les voyelles longues entre elles ; les brèves quant à elles, n'ont pas été différenciées par les auditeurs marocains. Nous observons que l'utilisation d'une durée longue des stimuli (correspondant à 200 ms) a davantage facilité l'identification des catégories vocaliques, tandis que les stimuli brefs (i.e., 100 ms) n'a pas aidé les auditeurs marocains. Les trois voyelles longues /i: a: u:/ ont été majoritairement catégorisées, avec les stimuli longs, tandis que les brèves /ɪ ʊ u/ ne l'étaient pas, ni avec les stimuli longs, ni avec les brefs. Comment peut-on alors interpréter les résultats des auditeurs marocains ? Ce résultat est difficile à expliquer lorsqu'on compare les résultats des **prototypes** identifiés par les auditeurs marocains et les résultats de catégorisation opérée par les mêmes auditeurs. Lors de la première tâche, les auditeurs arrivaient facilement à identifier trois **prototypes** significativement différents pour la voyelle /ə/ ; avec la seconde, les confusions de

catégorisation ont affecté leurs choix. Est-ce que les trois voyelles brèves identifiées avec exactitude par les auditeurs marocains dans l'expérience d'identification des **prototypes** (voir § 5.3.2) n'est qu'un artefact de tâche ? Les auditeurs étaient peut-être « obligés » d'identifier un **prototype** pour chaque voyelle, tandis que pour la tâche de catégorisation, ils avaient le choix entre trois différents « représentants » d'une « possible même catégorie vocalique » et la case « aucun ne correspond » ? Peut-on supposer que la tâche d'identification des **prototypes** rend davantage compte des « variations phonétiques » existant entre les stimuli proposés (i.e., un phénomène *one-to-many*, tandis que la tâche de catégorisation rend davantage compte de la « représentation » de ces stimuli pour chaque auditeur (i.e., un phénomène *many-to-one*) ? nous reviendrons sur cette question dans la discussion générale.

Intéressons-nous maintenant aux résultats de catégorisation de l'espace vocalique en AJ en **statique** vs **dynamique**.

5.4.3. *En Arabe Jordanien*

5.4.3.1. **Résultats en statique (V Target)**

Nous présentons dans le Tableau 55 les pourcentages de réponses des auditeurs jordaniens à l'expérience de catégorisation de l'espace vocalique avec des stimuli **statiques** longs. D'une façon générale, les auditeurs jordaniens étaient perturbés par cette expérience. Nous pouvons en effet voir dans ce tableau les confusions des auditeurs dans la catégorisation de toutes les voyelles. Seules les quatre voyelles longues /i: a: o: u:/ ont été bien identifiées par les auditeurs jordaniens avec des pourcentages supérieures à 50%. Nous pouvons néanmoins identifier trois grandes « zones de confusion » dans les choix des auditeurs jordaniens : la zone des voyelles /i: i e:/, la zone /a a:/ et la zone /o: u u:/ . Aucune voyelle brève n'a été identifiée par eux. Afin de mieux visualiser les résultats, nous les présentons sur la Figure 172 en fonction des sommes des scores des bonnes réponses.

RÉSULTATS EN PERCEPTION DES VOYELLES

Tableau 55 : Pourcentages de catégorisation des 2800 stimuli **statiques** (avec une durée de 200 ms) en AJ. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	ii	i	e:	a:	a	o:	u	u:	aucun	Total
1	250	2350	3100	54%	30%	4%	4%				2%	6%	100%
2	250	2150	3100	52%	20%	6%	6%			2%	6%	8%	100%
3	250	1950	2900	32%	18%	16%	4%		4%		6%	20%	100%
4	250	1750	2900	18%	8%	12%	2%	2%	2%	2%	18%	36%	100%
5	250	1500	2500						14%	4%	54%	28%	100%
6	250	1250	2500			2%	2%		14%		54%	28%	100%
7	250	1050	2300		2%				16%		68%	14%	100%
8	250	850	2300						16%		64%	20%	100%
9	250	650	2200						14%		78%	8%	100%
10	350	2350	3100	36%	26%	28%	2%		2%			6%	100%
11	350	2150	3100	26%	24%	28%	8%	2%	2%	2%	2%	6%	100%
12	350	1950	2900	24%	14%	26%	2%	2%	8%	2%	6%	16%	100%
13	350	1750	2900	12%	8%	12%			12%		26%	30%	100%
14	350	1500	2500		4%	8%			22%	6%	36%	24%	100%
15	350	1250	2500		2%	4%			32%		38%	24%	100%
16	350	1050	2300						50%	8%	30%	12%	100%
17	350	850	2300	2%			2%		50%	2%	38%	6%	100%
18	350	650	2200					2%	58%		30%	10%	100%
19	450	2150	3100	8%	26%	32%	12%			2%	2%	18%	100%
20	450	1950	2900	4%	12%	34%	18%		12%	2%		18%	100%
21	450	1750	2900		6%	26%	12%		12%	2%	20%	22%	100%
22	450	1500	2500	4%	4%	12%	20%	2%	26%	2%	10%	20%	100%
23	450	1250	2500		2%	4%	14%	2%	28%	2%	20%	28%	100%
24	450	1050	2300			2%	12%		58%		10%	18%	100%
25	450	850	2300	4%	2%		8%		58%	2%	14%	12%	100%
26	450	650	2200		2%		12%		48%	2%	24%	12%	100%
27	550	2150	3100	6%	18%	26%	32%	2%	2%	2%	2%	12%	100%
28	550	1950	2900	2%	6%	28%	32%	2%	6%	2%	2%	22%	100%

N° Stimuli	F ₁	F ₂	F ₃	ii	i	e:	a:	a	o:	u	u:	aucun	Total
29	550	1750	2900		4%	16%	44%		2%		4%	30%	100%
30	550	1500	2500		2%	6%	48%		8%	4%	8%	24%	100%
31	550	1250	2500		4%		58%		18%	4%	2%	14%	100%
32	550	1050	2300				56%	4%	20%		2%	18%	100%
33	550	850	2300		2%		34%	4%	36%		6%	18%	100%
34	550	650	2200				48%		20%	6%	2%	24%	100%
35	650	1950	2900		4%	14%	38%	6%	2%			36%	100%
36	650	1750	2900		2%		66%	2%	4%	4%	2%	22%	100%
37	650	1500	2500	2%	2%		88%			2%		6%	100%
38	650	1250	2500				86%	2%	2%		4%	6%	100%
39	650	1050	2300	2%	4%		80%				4%	10%	100%
40	650	850	2300				68%	4%	6%	2%	4%	16%	100%
41	750	1950	2900		2%	2%	56%	2%	2%	2%	6%	28%	100%
42	750	1750	2900		2%		58%		2%		6%	32%	100%
43	750	1500	2500	2%			82%	4%	4%	2%	2%	4%	100%
44	750	1250	2500		2%		82%				6%	10%	100%
45	750	1050	2300		2%	2%	86%	2%			2%	6%	100%
46	750	850	2300		2%	2%	74%			4%		18%	100%
47	250	1950	2300	28%	16%	14%	2%		2%		10%	28%	100%
48	350	1950	2300	16%	26%	28%	2%		10%		4%	14%	100%
49	450	1750	2300		8%	28%	24%	2%	10%	2%	12%	14%	100%
50	550	1750	2300		4%	20%	36%	2%	4%	2%	6%	26%	100%
51	650	1500	2300		2%		76%	2%	2%	2%	2%	14%	100%
52	750	1500	2300		4%		78%	2%	2%			14%	100%
53	250	850	2700					2%	26%	4%	58%	10%	100%
54	350	850	2700	2%	2%				44%	2%	44%	6%	100%
55	450	850	2700		2%	4%	10%	2%	48%		20%	14%	100%
56	550	850	2700	4%	2%		30%		38%		4%	22%	100%

L'examen attentif de la Figure 172 révèle que seules les trois voyelles /a: o: u:/ ont été majoritairement catégorisées par au moins sept auditeurs jordaniens sur dix. La voyelle /a:/ a été catégorisée à partir de dix stimuli différents, /o:/ à partir d'un seul stimulus et /u:/ à partir de trois. La voyelle /i:/ a été confondue à deux reprises soit avec /i/, soit avec /i e:/. La voyelle /e:/ n'a jamais été catégorisée majoritairement par les auditeurs jordaniens, mais a été confondue. Les trois voyelles brèves ont été mal identifiées à cause de leurs rapprochements avec les voyelles longues correspondantes. Si l'on compare les pourcentages d'identification dans le Tableau 55 et la catégorisation présentée sur la Figure 172, nous observons que seule la voyelle /i/ a été identifiée par les auditeurs jordaniens. Les deux autres voyelles brèves /a u/ ont été ignorées dans la catégorisation car confondues avec les voyelles longues correspondantes. Ces résultats ont permis d'établir les frontières entre les catégories vocaliques longues uniquement.

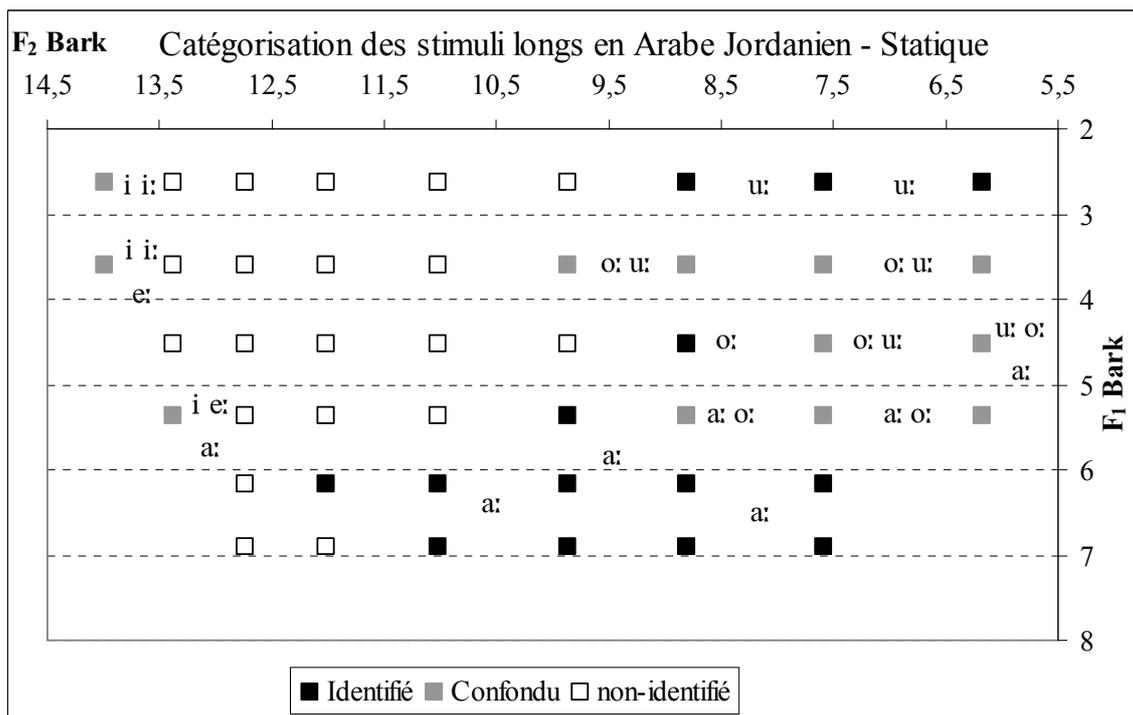


Figure 172 : Catégorisation de l'espace vocalique en AJ en **statique** avec une durée de 200 ms. « Identifié » même catégorie, ≥ 7 , « Confondu » plusieurs catégories ≥ 7 , « non-identifié » = une ou plusieurs catégories ≤ 6 .

Intéressons-nous maintenant aux pourcentages de réponses des auditeurs jordaniens à partir des stimuli **statiques** courts (voir Tableau 56).

Tableau 56 : Pourcentages de catégorisation des 2800 stimuli **statiques** (avec une durée de 100 ms) en AJ. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	i:	i	e:	a:	a	o:	u	u:	aucun	Total
57	250	2350	3100	24%	42%	2%	4%	6%	2%	2%	2%	16%	100%
58	250	2150	3100	24%	46%	10%	2%	8%	2%	2%	8%	8%	100%
59	250	1950	2900	14%	30%	8%	2%	2%	6%	8%	6%	26%	100%
60	250	1750	2900	14%	10%	6%	2%	4%	6%	14%	14%	30%	100%
61	250	1500	2500	2%	2%	2%	2%	4%	8%	22%	26%	36%	100%
62	250	1250	2500	2%	6%	2%	2%	4%	10%	16%	38%	30%	100%
63	250	1050	2300	6%	2%	2%	2%	4%	12%	22%	40%	16%	100%
64	250	850	2300	2%	2%	2%	4%	4%	16%	26%	38%	14%	100%
65	250	650	2200	2%	2%	2%	2%	2%	18%	26%	32%	20%	100%
66	350	2350	3100	8%	34%	20%	2%	8%	2%	4%	4%	20%	100%
67	350	2150	3100	8%	46%	8%	2%	6%	10%	2%	4%	18%	100%
68	350	1950	2900	4%	40%	2%	2%	8%	2%	2%	6%	34%	100%
69	350	1750	2900	4%	22%	2%	10%	8%	4%	8%	10%	34%	100%
70	350	1500	2500	2%	2%	4%	6%	2%	16%	28%	4%	38%	100%
71	350	1250	2500	2%	2%	2%	2%	2%	20%	28%	26%	24%	100%
72	350	1050	2300	2%	2%	2%	2%	2%	18%	32%	32%	16%	100%
73	350	850	2300	2%	2%	2%	2%	2%	30%	32%	12%	24%	100%
74	350	650	2200	2%	2%	2%	2%	2%	30%	24%	20%	24%	100%
75	450	2150	3100	36%	36%	4%	20%	4%	4%	2%	2%	32%	100%
76	450	1950	2900	2%	36%	4%	14%	6%	4%	6%	2%	26%	100%
77	450	1750	2900	4%	20%	8%	8%	4%	6%	20%	2%	30%	100%
78	450	1500	2500	10%	10%	2%	8%	16%	10%	14%	8%	34%	100%
79	450	1250	2500	2%	2%	2%	2%	4%	28%	34%	8%	24%	100%
80	450	1050	2300	2%	2%	2%	4%	2%	26%	30%	6%	30%	100%
81	450	850	2300	2%	2%	4%	14%	6%	20%	24%	8%	22%	100%
82	450	650	2200	4%	4%	4%	4%	6%	22%	28%	2%	34%	100%
83	550	2150	3100	4%	28%	2%	18%	8%	2%	2%	2%	36%	100%
84	550	1950	2900	32%	32%	2%	22%	12%	4%	4%	2%	22%	100%

N° Stimuli	F ₁	F ₂	F ₃	i:	i	e:	a:	a	o:	u	u:	aucun	Total
85	550	1750	2900	2%	22%	2%	20%	16%	2%	4%	4%	28%	100%
86	550	1500	2500	4%	2%	4%	26%	22%	2%	10%	6%	30%	100%
87	550	1250	2500	2%	2%	4%	16%	24%	12%	16%	2%	26%	100%
88	550	1050	2300	4%	4%	4%	14%	12%	8%	18%	2%	40%	100%
89	550	850	2300	6%	6%	8%	12%	14%	14%	14%	2%	40%	100%
90	550	650	2200	10%	10%	10%	10%	10%	8%	14%	2%	40%	100%
91	650	1950	2900	24%	24%	10%	26%	8%	4%	2%	2%	24%	100%
92	650	1750	2900	10%	14%	2%	26%	28%	2%	4%	2%	28%	100%
93	650	1500	2500	14%	14%	2%	30%	24%	2%	2%	2%	26%	100%
94	650	1250	2500	2%	2%	2%	30%	28%	2%	2%	6%	30%	100%
95	650	1050	2300	2%	2%	2%	36%	16%	2%	6%	2%	38%	100%
96	650	850	2300	2%	2%	2%	28%	22%	4%	2%	2%	38%	100%
97	750	1950	2900	12%	12%	2%	10%	20%	4%	4%	4%	52%	100%
98	750	1750	2900	8%	8%	2%	20%	22%	4%	4%	2%	38%	100%
99	750	1500	2500	6%	6%	2%	30%	24%	2%	4%	4%	28%	100%
100	750	1250	2500	6%	6%	2%	40%	24%	2%	2%	2%	26%	100%
101	750	1050	2300	6%	6%	2%	48%	18%	2%	2%	2%	26%	100%
102	750	850	2300	2%	2%	2%	26%	18%	4%	10%	2%	38%	100%
103	250	1950	2300	24%	24%	16%	6%	6%	4%	4%	12%	34%	100%
104	350	1950	2300	12%	34%	4%	4%	8%	6%	6%	2%	28%	100%
105	450	1750	2300	24%	24%	6%	20%	12%	2%	6%	6%	24%	100%
106	550	1750	2300	22%	4%	4%	22%	20%	4%	4%	4%	22%	100%
107	650	1500	2300	4%	4%	2%	34%	20%	4%	8%	2%	28%	100%
108	750	1500	2300	8%	8%	2%	36%	18%	2%	2%	2%	34%	100%
109	250	850	2700	2%	2%	2%	2%	2%	26%	24%	30%	14%	100%
110	350	850	2700	2%	2%	2%	4%	4%	12%	32%	26%	26%	100%
111	450	850	2700	4%	4%	4%	14%	6%	16%	22%	8%	30%	100%
112	550	850	2700	4%	4%	2%	10%	6%	20%	18%	4%	42%	100%

L'utilisation des stimuli brefs a perturbé les auditeurs jordaniens, comme les marocains ; ceci est visible directement à partir de leurs réponses. Afin de mieux visualiser les résultats de catégorisation, nous les présentons sur la Figure 173. La plupart des stimuli confondus par les auditeurs jordaniens correspondent soit aux voyelles /i a a:/, soit à /o: u u:/. Les deux voyelles longues /i: e:/ n'ont jamais été choisies.

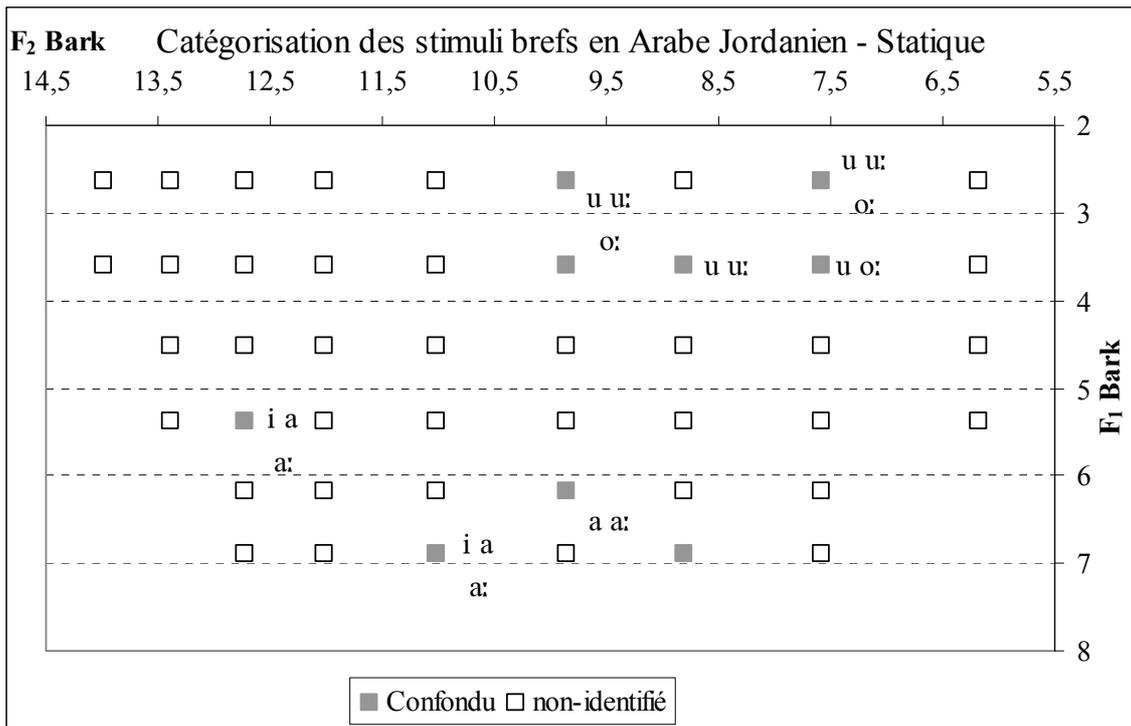


Figure 173 : Catégorisation de l'espace vocalique en AJ en **statique** avec une durée de 100 ms. « Confondu » plusieurs catégories ≥ 7 , « non-identifié » = une ou plusieurs catégories ≤ 6 .

Comme nous l'avons vu précédemment, la **dynamique** a aidé énormément les auditeurs jordaniens dans la tâche d'identification des **prototypes** vocaliques (voir § 5.3.3.2). Nous présentons dans la partie qui suit les résultats de catégorisation de l'espace vocalique en AJ en **dynamique**.

5.4.3.2. Résultats en dynamique (CV Target)

Dans le Tableau 57, nous présentons les pourcentages de catégorisation des stimuli en **dynamique** en AJ.

Tableau 57 : Pourcentages de catégorisation des 2800 stimuli **dynamiques** (avec une durée de 200 ms) en AJ. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	i:	i	e:	a:	a	o:	u	u:	aucun	Total
1	250	2350	3100	50%	18%	10%				10%	2%	10%	100%
2	250	2150	3100	46%	28%	12%				6%	4%	4%	100%
3	250	1950	2900	40%	18%	14%			4%	6%	4%	14%	100%
4	250	1750	2900	8%	14%	10%			6%	4%	34%	24%	100%
5	250	1500	2500	6%	2%	2%	2%	4%	18%	2%	46%	18%	100%
6	250	1250	2500	4%			2%	2%	14%		62%	18%	100%
7	250	1050	2300				4%	2%	18%	4%	64%	8%	100%
8	250	850	2300			4%		6%	30%	2%	56%	2%	100%
9	250	650	2200				4%		8%	4%	54%	30%	100%
10	350	2350	3100	22%	28%	28%	8%		2%	6%		6%	100%
11	350	2150	3100	16%	22%	34%	6%			12%		10%	100%
12	350	1950	2900	10%	6%	28%	2%		2%	10%	12%	30%	100%
13	350	1750	2900		10%	20%			8%	8%	28%	26%	100%
14	350	1500	2500		6%	8%		4%	18%	4%	42%	18%	100%
15	350	1250	2500	8%		2%	2%		24%	4%	24%	36%	100%
16	350	1050	2300				4%	2%	26%	8%	40%	20%	100%
17	350	850	2300				6%	2%	64%	12%	8%	8%	100%
18	350	650	2200					2%	54%		18%	26%	100%
19	450	2150	3100	6%	18%	38%	6%		2%	4%	4%	22%	100%
20	450	1950	2900		14%	42%	4%		2%	10%	2%	26%	100%
21	450	1750	2900	4%	2%	32%	8%	4%	12%	6%	2%	30%	100%
22	450	1500	2500		2%	18%	22%	8%	24%	6%	6%	14%	100%
23	450	1250	2500		2%	10%	12%	2%	38%	4%	2%	30%	100%
24	450	1050	2300			4%	22%	2%	30%	12%	12%	18%	100%
25	450	850	2300		2%		6%		60%	4%	10%	18%	100%
26	450	650	2200				4%		54%	4%	10%	28%	100%
27	550	2150	3100	4%	12%	28%	8%			8%	2%	38%	100%
28	550	1950	2900	4%	8%	32%	20%		4%	6%		26%	100%

N° Stimuli	F ₁	F ₂	F ₃	i:	i	e:	a:	a	o:	u	u:	aucun	Total
29	550	1750	2900	6%		20%	26%			14%	6%	4%	100%
30	550	1500	2500		2%	6%	52%	2%	10%	2%	6%	20%	100%
31	550	1250	2500			2%	48%	2%	26%	2%	2%	18%	100%
32	550	1050	2300				50%	2%	16%	6%	2%	24%	100%
33	550	850	2300	2%		2%	44%		30%	4%	4%	14%	100%
34	550	650	2200				36%		24%	6%	4%	30%	100%
35	650	1950	2900		6%	24%	42%			6%		22%	100%
36	650	1750	2900		4%	10%	68%		2%	4%	2%	10%	100%
37	650	1500	2500	2%	4%		30%	22%		14%	4%	24%	100%
38	650	1250	2500		2%		78%	2%	2%	2%	2%	12%	100%
39	650	1050	2300		4%		80%		2%		2%	12%	100%
40	650	850	2300		2%		66%	2%	4%	2%		24%	100%
41	750	1950	2900		2%	6%	64%		4%	2%	2%	20%	100%
42	750	1750	2900			2%	86%					12%	100%
43	750	1500	2500				78%	2%	2%	2%	2%	14%	100%
44	750	1250	2500				82%			4%	2%	12%	100%
45	750	1050	2300				76%		2%	2%	4%	16%	100%
46	750	850	2300		2%		72%	2%		4%		20%	100%
47	250	1950	2300	30%	24%	8%		2%	2%	4%	8%	22%	100%
48	350	1950	2300	4%	18%	38%	6%	2%	2%	8%	2%	20%	100%
49	450	1750	2300		6%	34%	14%	2%	14%	8%	4%	18%	100%
50	550	1750	2300		8%	22%	24%	2%	12%	2%	6%	24%	100%
51	650	1500	2300			2%	74%			6%		18%	100%
52	750	1500	2300		2%	2%	78%	2%				16%	100%
53	250	850	2700						18%		78%	4%	100%
54	350	850	2700		2%				62%	8%	16%	12%	100%
55	450	850	2700	6%		6%	8%		38%	4%	18%	20%	100%
56	550	850	2700				42%	6%	24%	4%	4%	20%	100%

dynamiques, /a:/ observe le même nombre de stimuli identifiés mais change de position et /o:/ obtient plus de catégories majoritairement identifiées. Dans les deux cas (i.e., avec les stimuli **statiques** et **dynamiques**), la voyelle /a:/ n'a pas été catégorisée en fonction de tous les stimuli « ouverts » (i.e., avec un F₁ à 6 et à 7 Bark). Les deux derniers stimuli ouverts présentant un F₂ à 12,7 Bark ont été confondus avec la voyelle /e:/. Phonétiquement, ces deux stimuli représentent la voyelle [æ]. Dans le contexte dental choisi [da:r], les auditeurs jordaniens ont choisi une voyelle /a:/ centrale et postérieure à cause de la présence de [r].

Nous nous intéressons dans le Tableau 58, aux pourcentages de réponses des auditeurs jordaniens, pour la catégorisation avec les stimuli **dynamiques** courts (i.e., à 100 ms).

Nous observons que seuls quatre stimuli **dynamiques** ont été identifiés au moins à 50% de bonnes réponses par les auditeurs jordaniens. Ces quatre stimuli ont été catégorisés soit comme /i/, soit comme /u/. Autrement dit, les auditeurs jordaniens étaient conscients de la variation de durée pour ne choisir que ces deux voyelles brèves. À l'inverse des réponses avec les stimuli **statiques**, la grande majorité des « bonnes » catégorisations, même si elles ne dépassent pas les 50%, sont obtenues pour les voyelles brèves. Seule la voyelle longue /a:/ a été sélectionnée comme « représentante » de la catégorie des voyelles « ouvertes ».

Tableau 58 : Pourcentages de catégorisation des 2800 stimuli **dynamiques** (avec une durée de 100 ms) en AJ. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	i̇	i	e:	a:	a	o:	u	u:	aucun	Total
57	250	2350	3100	10%	50%	6%		2%		12%	2%	18%	100%
58	250	2150	3100	2%	44%	12%	2%	2%	2%	12%		24%	100%
59	250	1950	2900	16%	34%	2%		4%		14%	4%	26%	100%
60	250	1750	2900	4%	16%			6%		22%	20%	32%	100%
61	250	1500	2500					16%	34%	32%	18%	100%	100%
62	250	1250	2500		2%			2%	8%	40%	36%	12%	100%
63	250	1050	2300		2%		2%	2%	12%	34%	36%	12%	100%
64	250	850	2300		2%		4%	4%	2%	40%	32%	16%	100%
65	250	650	2200		4%			2%	10%	36%	28%	20%	100%
66	350	2350	3100	10%	38%	2%	2%	8%	6%	6%	2%	26%	100%
67	350	2150	3100	12%	32%	12%	2%		6%	12%		24%	100%
68	350	1950	2900	4%	30%	6%		10%	2%	20%	8%	20%	100%
69	350	1750	2900	8%	24%		2%	6%	4%	22%	14%	20%	100%
70	350	1500	2500		8%			2%	22%	38%	16%	14%	100%
71	350	1250	2500					2%	16%	50%	22%	10%	100%
72	350	1050	2300		6%		2%	6%	10%	56%	6%	14%	100%
73	350	850	2300					2%	2%	50%	16%	30%	100%
74	350	650	2200		8%	2%		6%	10%	34%	12%	28%	100%
75	450	2150	3100		18%	18%	2%	4%	2%	18%		38%	100%
76	450	1950	2900	6%	30%	8%	2%	6%	6%	16%	6%	26%	100%
77	450	1750	2900	4%	24%	4%	6%	10%	8%	20%	6%	18%	100%
78	450	1500	2500		4%	8%		10%	12%	40%	10%	16%	100%
79	450	1250	2500		2%	4%	8%	12%	8%	30%	18%	18%	100%
80	450	1050	2300		4%		8%	6%	24%	40%	2%	16%	100%
81	450	850	2300		2%		8%	2%	14%	38%	8%	28%	100%
82	450	650	2200	2%			4%	2%	4%	36%	12%	40%	100%
83	550	2150	3100		24%	14%	2%	8%	8%	16%		36%	100%
84	550	1950	2900		26%	8%		2%	4%	18%	2%	40%	100%
85	550	1750	2900	8%	16%	6%	10%	20%	2%	18%		20%	100%
86	550	1500	2500			4%	20%	20%	14%	18%		24%	100%
87	550	1250	2500				10%	28%	16%	24%	2%	20%	100%
88	550	1050	2300		6%		16%	20%	10%	26%	2%	20%	100%
89	550	850	2300			2%	12%	8%	14%	44%	8%	12%	100%
90	550	650	2200				20%	14%	8%	28%	4%	26%	100%
91	650	1950	2900		12%	6%	12%	30%	2%	10%	2%	26%	100%
92	650	1750	2900		6%		18%	28%	4%	4%	6%	34%	100%
93	650	1500	2500		2%		26%	22%	4%	16%	2%	28%	100%
94	650	1250	2500		2%		24%	34%	2%	16%		22%	100%
95	650	1050	2300	2%			24%	22%	6%	14%	6%	26%	100%
96	650	850	2300				20%	26%	2%	20%		32%	100%
97	750	1950	2900	2%	6%	2%	12%	38%		8%	4%	28%	100%
98	750	1750	2900				16%	32%	4%	10%	4%	34%	100%
99	750	1500	2500		10%		26%	28%		6%	2%	28%	100%
100	750	1250	2500		4%		34%	24%		6%	2%	30%	100%
101	750	1050	2300		4%		32%	28%	2%	10%	2%	22%	100%
102	750	850	2300				26%	28%	2%	14%	4%	26%	100%
103	250	1950	2300	6%	20%	6%	2%			14%	12%	40%	100%
104	350	1950	2300	10%	34%	12%		4%		14%	2%	24%	100%
105	450	1750	2300	12%	14%	12%		8%	4%	18%	4%	28%	100%
106	550	1750	2300		8%	6%	10%	20%	6%	22%		28%	100%
107	650	1500	2300		2%		34%	32%	4%	12%		16%	100%
108	750	1500	2300		4%		36%	26%	4%	4%	2%	24%	100%
109	250	850	2700					4%	14%	18%	48%	16%	100%
110	350	850	2700	2%	4%		4%	4%	18%	40%	12%	16%	100%
111	450	850	2700				4%		24%	48%	10%	14%	100%
112	550	850	2700				4%	16%	20%	38%	6%	16%	100%

La Figure 175 montre les résultats de la catégorisation des stimuli **dynamiques** brefs. Aucun stimulus n'a été catégorisé majoritairement par les auditeurs jordaniens. La plupart des stimuli avec une position centrale et postérieure, /a a: o: u u:/ ont été confondus.

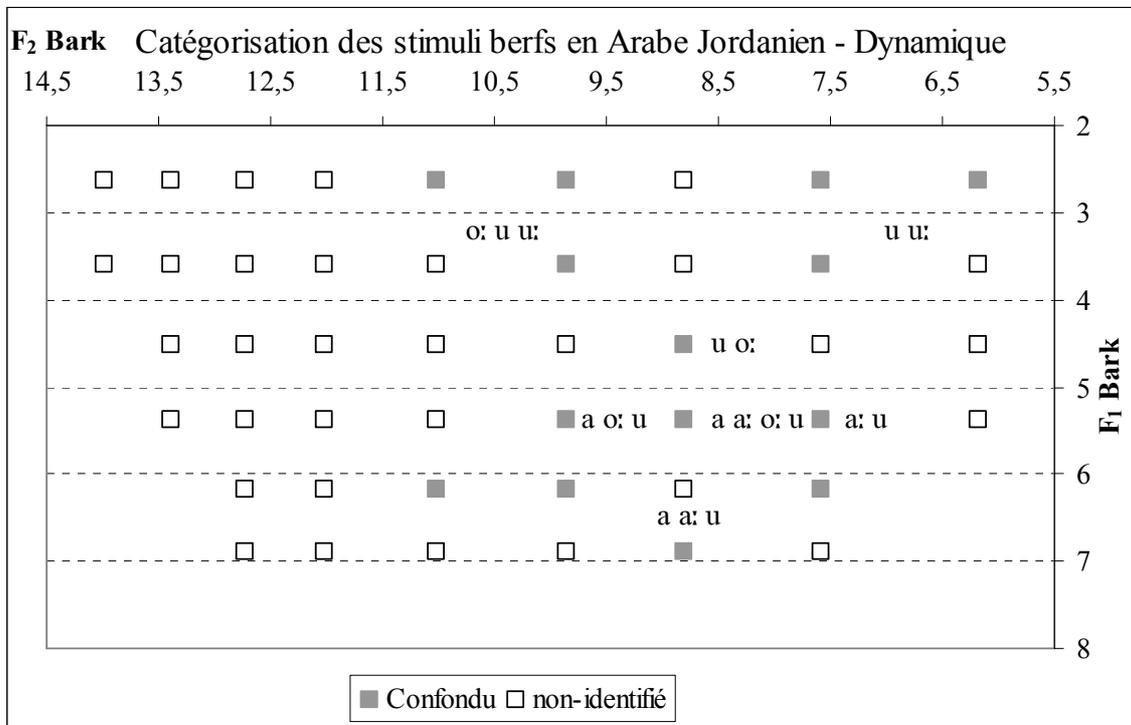


Figure 175 : Catégorisation de l'espace vocalique en AJ en **dynamique** avec une durée de 100 ms. « Confondu » plusieurs catégories ≥ 7 , « non-identifié » = une ou plusieurs catégories ≤ 6 .

Globalement, les auditeurs jordaniens sont capables de catégoriser les stimuli **statiques** ou **dynamiques** en fonction des voyelles de leur(s) système(s). Comme en AM, les résultats de catégorisation de l'espace vocalique en AJ montrent que cette tâche est « plus » facile avec les stimuli longs (avec une durée de 200 ms) qu'avec les stimuli brefs (avec une durée à 100 ms). Les auditeurs jordaniens ont expliqué que ce sont les différences de timbres qui les ont aidées dans leurs catégorisations et non les différences de durée. Les trois voyelles longues /a: o: u:/ ont été majoritairement catégorisées, avec les stimuli longs, les autres voyelles, /i: i e: a u/ ont été confondues. Les voyelles brèves n'ont jamais été identifiées, car confondues avec leurs correspondantes longues.

Nous nous intéressons dans la partie qui suit aux résultats de catégorisation de l'espace vocalique en FR dans les deux modalités **statique** vs **dynamique**.

5.4.4. *En Français*

5.4.4.1. **Résultats en statique (V Target)**

Nous présentons dans le Tableau 59 les pourcentages de réponses des auditeurs français. L'examen des résultats révèle que les auditeurs français sont plus précis dans leurs catégorisations par rapport aux auditeurs marocains ou jordaniens. En effet, nous observons moins de confusions en FR par rapport aux deux autres langues. Toutes les voyelles en FR ont été catégorisées avec plus de 50% de bonnes réponses. Les quelques confusions observées sont faites entre des voyelles proches sur l'un des deux axes, F_1 ou F_2 .

Les résultats de la catégorisation des auditeurs français, en fonction des scores des bonnes réponses (voir Figure 176), montrent que les voyelles /i e ε a o u y ø œ/ ont été majoritairement catégorisées par au moins sept auditeurs français sur dix. Les deux voyelles /ɑ ɔ/ ont été confondues soit entre elles, soit avec /a o/, respectivement. Nous pensons alors que les confusions observées pour la voyelle /ɑ/ peuvent expliquer que le statut phonologique de celle-ci a bel et bien changé pour les auditeurs mêmes. Nous observons également que les confusions au niveau des voyelles médianes antérieures /e ε ø œ/, et plus particulièrement pour les stimuli dont le F_1 est autour de 4,5 Bark, donnent une bonne indication de la difficulté des auditeurs français à différencier celles-ci. Les confusions observées en FR permettent d'identifier les frontières entre les voyelles.

Comme nous pouvons l'observer, le nombre de stimuli identifiant une catégorie vocalique quelconque ne dépasse pas trois. Les auditeurs français sont plus précis dans leurs choix que les auditeurs marocains ou jordaniens, car la dispersion intra-catégorie vocalique est réduite en FR par rapport aux deux autres langues.

Tableau 59 : Pourcentages de catégorisation des 2800 stimuli **statiques** (avec une durée de 200 ms) en FR. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	F ₁	F ₂	F ₃	i	e	ε	a	α	ɔ	o	u	y	θ	œ	aucun	Total
1	250	2350	3100	96%	4%											100%
2	250	2150	3100	90%	10%											100%
3	250	1950	2900	48%	18%	2%					24%				8%	100%
4	250	1750	2900	36%	10%	4%					30%				20%	100%
5	250	1500	2500		6%							52%	26%		16%	100%
6	250	1250	2500							2%	12%	22%	48%		16%	100%
7	250	1050	2300							2%	40%	4%	24%		30%	100%
8	250	850	2300							2%	84%	2%	10%		2%	100%
9	250	650	2200							2%	94%				4%	100%
10	350	2350	3100	4%	88%	6%									2%	100%
11	350	2150	3100		94%	4%									2%	100%
12	350	1950	2900		56%	10%					8%		18%	4%	4%	100%
13	350	1750	2900		36%	10%							46%	2%	6%	100%
14	350	1500	2500				2%						86%	12%		100%
15	350	1250	2500							2%			92%	6%		100%
16	350	1050	2300							2%	16%		60%	12%	10%	100%
17	350	850	2300					16%	56%	4%			4%	12%	8%	100%
18	350	650	2200					6%	92%					2%		100%
19	450	2150	3100		78%	22%							20%		10%	100%
20	450	1950	2900		42%	28%							36%	18%	14%	100%
21	450	1750	2900	2%	18%	12%							58%	40%	2%	100%
22	450	1500	2500										54%	40%	6%	100%
23	450	1250	2500							8%	28%		16%	26%	22%	100%
24	450	1050	2300						8%	32%	58%		6%	4%	6%	100%
25	450	850	2300					2%	44%	42%			2%	10%	10%	100%
26	450	650	2200												6%	100%
27	550	2150	3100		14%	80%									6%	100%
28	550	1950	2900		12%	66%									12%	100%

N° Stimuli	F ₁	F ₂	F ₃	i	e	ε	a	α	ɔ	o	u	y	θ	œ	aucun	Total
29	550	1750	2900			22%							10%	50%	18%	100%
30	550	1500	2500										14%	76%	10%	100%
31	550	1250	2500						8%				20%	58%	4%	100%
32	550	1050	2300						6%	18%	26%	12%		20%	18%	100%
33	550	850	2300						18%	52%	12%			4%	14%	100%
34	550	650	2200						2%	18%	34%	14%		6%	26%	100%
35	650	1950	2900		78%									6%	8%	100%
36	650	1750	2900		2%	20%							10%	44%	24%	100%
37	650	1500	2500			2%		4%					14%	64%	16%	100%
38	650	1250	2500					16%	20%	10%				36%	18%	100%
39	650	1050	2300					40%	40%	14%				2%	4%	100%
40	650	850	2300					36%	38%	10%	4%				12%	100%
41	750	1950	2900		4%	56%	4%	2%						4%	30%	100%
42	750	1750	2900		4%	8%	4%	8%					4%	44%	28%	100%
43	750	1500	2500					58%	34%					6%	2%	100%
44	750	1250	2500					64%	30%				2%		4%	100%
45	750	1050	2300					60%	40%							100%
46	750	850	2300					32%	48%	4%	4%				12%	100%
47	250	1950	2300	4%	10%	2%						80%			4%	100%
48	350	1950	2300		44%	8%						24%			10%	100%
49	450	1750	2300		16%	4%						2%	38%	30%	10%	100%
50	550	1750	2300			26%							18%	22%	34%	100%
51	650	1500	2300			2%		4%					12%	62%	20%	100%
52	750	1500	2300			4%	8%	20%		4%			8%	30%	26%	100%
53	250	850	2700						2%		90%		8%			100%
54	350	850	2700						10%	78%			4%	6%	2%	100%
55	450	850	2700					2%	32%	60%					6%	100%
56	550	850	2700				4%	22%	42%	12%				4%	16%	100%

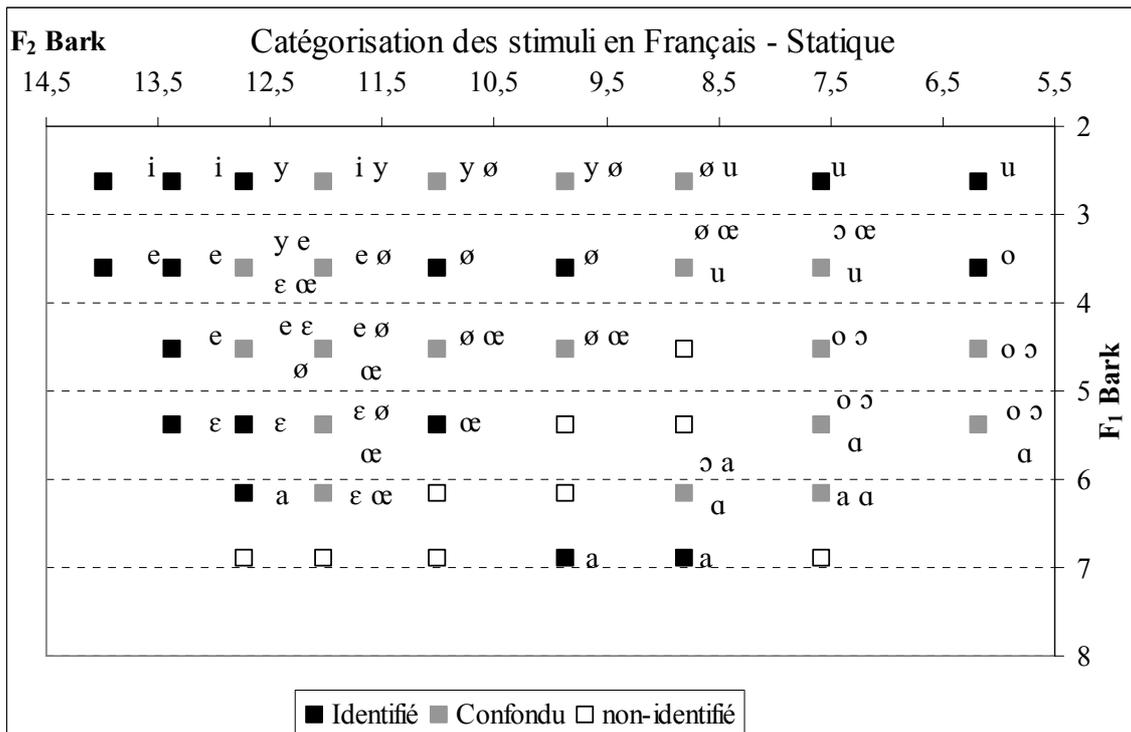


Figure 176 : Catégorisation de l'espace vocalique en FR en **statique** avec une durée de 200 ms. « Identifié » même catégorie, ≥ 7 , « Confondu » plusieurs catégories ≥ 7 , « non-identifié » = une ou plusieurs catégories ≤ 6 .

Nous présentons dans la partie qui suit les résultats de catégorisation de l'espace vocalique en FR en **dynamique**.

5.4.4.2. Résultats en dynamique (CV Target)

Nous présentons dans le Tableau 60 les pourcentages de catégorisation des stimuli **dynamiques** en FR. L'examen des résultats montre que les auditeurs français sont plus précis dans leurs catégorisations en utilisant les stimuli **dynamiques**. En effet, le pourcentage de bonnes réponses avoisine les 98% pour certaines voyelles. D'une manière générale, les auditeurs français arrivent à bien identifier les stimuli en les plaçant dans des catégories distinctes. Toutefois, nous observons des confusions pour les stimuli représentant les voyelles médianes.

RÉSULTATS EN PERCEPTION DES VOYELLES

Tableau 60 : Pourcentages de catégorisation des 2800 stimuli **dynamiques** (avec une durée de 200 ms) en FR. Le fond gris marque les cases dont le choix est égal à 0%. En gras sont marquées les réponses dont la catégorisation est \geq à 50%.

N° Stimuli	Stimuli 1-17										Stimuli 18-28										Total											
	F ₁	F ₂	F ₃	i	e	ε	a	α	o	u	y	ø	œ	aucun	Total	F ₁	F ₂	F ₃	i	e		ε	a	α	o	u	y	ø	œ	aucun	Total	
1	250	2350	3100	98%	2%										100%	350	650	850	2300										100%			
2	250	2150	3100	92%		2%				4%				2%	100%	350	1750	2500											2%	100%		
3	250	1950	2900	44%					48%					8%	100%	350	1500	2500											8%	100%		
4	250	1750	2900	26%	2%				56%					14%	100%	350	1050	2300											10%	100%		
5	250	1500	2500	6%					68%	16%				10%	100%	350	850	2300											4%	100%		
6	250	1250	2500						8%	34%	24%			28%	100%	350	650	2200										12%	20%	100%		
7	250	1050	2300						44%	10%	24%			22%	100%	350	950	2900										4%	2%	100%		
8	250	850	2300						94%			2%		4%	100%	350	1750	2900										34%	26%	100%		
9	250	650	2200					6%	94%					2%	100%	350	1500	2500										6%	16%	100%		
10	350	2350	3100	4%	82%	12%								2%	100%	350	1250	2500										26%	14%	100%		
11	350	2150	3100	6%	78%	14%								2%	100%	350	1050	2300										32%	40%	100%		
12	350	1950	2900		34%	26%						18%		22%	100%	350	850	2300											4%	100%		
13	350	1750	2900		14%	10%			4%			64%	4%	4%	100%	350	1950	2900										2%	24%	100%		
14	350	1500	2500			2%						88%	10%	100%	350	1750	2900											4%	24%	100%		
15	350	1250	2500									96%	4%	100%	350	1500	2500											8%	30%	100%		
16	350	1050	2300					6%	2%	4%		62%	4%	22%	100%	350	1250	2500										56%	38%	100%		
17	350	850	2300					8%	68%	8%			12%	4%	100%	350	1050	2300										44%	54%	2%	100%	
18	350	650	2200					6%	84%	8%				2%	100%	350	850	2300										58%	30%	4%	100%	
19	450	2150	3100		68%	32%								20%	100%	450	1950	2300											8%	100%		
20	450	1950	2900		34%	34%						6%	6%	20%	100%	450	1750	2300											2%	2%	100%	
21	450	1750	2900		10%	14%						38%	20%	18%	100%	450	1500	2300											26%	16%	100%	
22	450	1500	2500									28%	64%	8%	100%	550	1750	2300										2%	36%	100%		
23	450	1250	2500					4%	4%	4%		40%	40%	8%	100%	550	1500	2300										4%	2%	2%	100%	
24	450	1050	2300					30%	24%	4%		6%	26%	10%	100%	550	1500	2300										12%	18%	16%	100%	
25	450	850	2300					8%	36%	54%			2%	100%	550	850	2700												4%	4%	100%	
26	450	650	2200					8%	22%	60%				10%	100%	450	850	2700											4%	74%	12%	100%
27	550	2150	3100		2%	94%								4%	100%	550	1950	2900											2%	8%	100%	
28	550	1950	2900		22%	62%							10%	6%	100%	550	1750	2900										2%	30%	40%	22%	100%

Nous présentons les résultats de catégorisation en fonction des scores de réponses des auditeurs français sur la Figure 177. Comme avec les stimuli statiques, les voyelles /i e ε a o u y ø œ/ ont été majoritairement identifiées par au moins sept auditeurs français sur dix. Les deux voyelles /ɑ ɔ/ ont été confondues avec des voyelles proches au niveau de F₁ ou de F₂. Tous les stimuli dont le F₁ est à 4,5 Bark (hormis le stimulus catégorisé comme /o/) ont été confondus dans leurs catégorisations : les auditeurs français ont choisi les voyelles médianes pour les catégoriser.

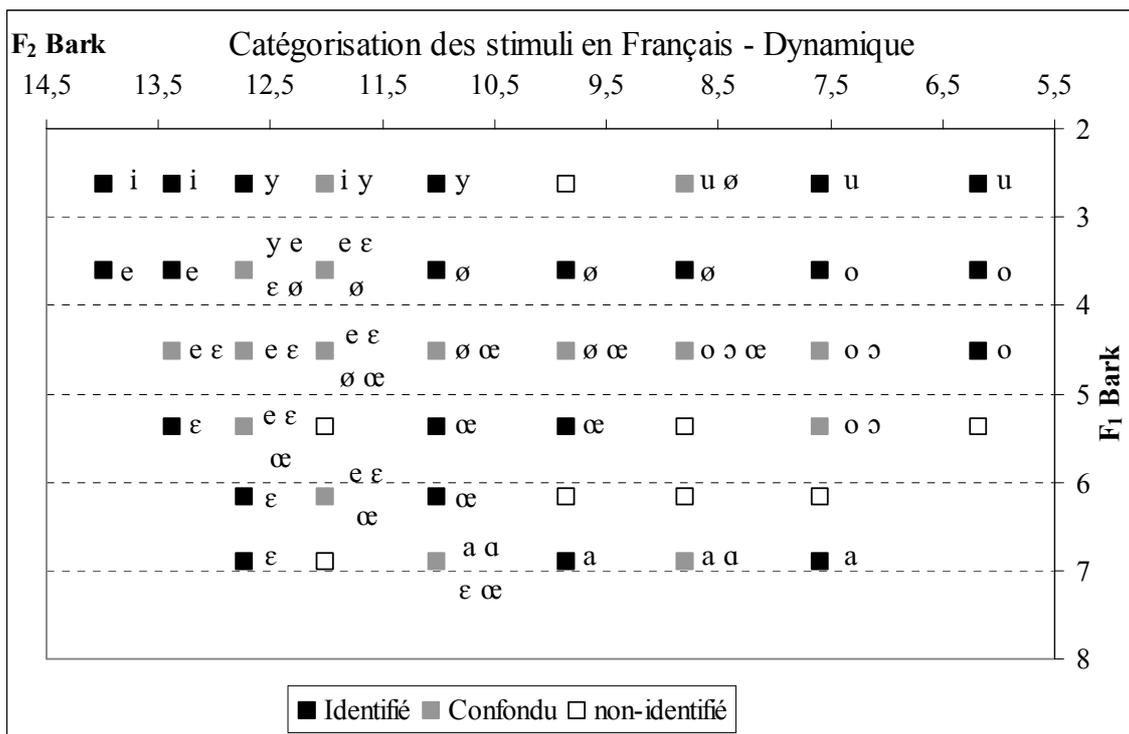


Figure 177 : Catégorisation de l'espace vocalique en FR en **dynamique** avec une durée de 200 ms. « Identifié » même catégorie, ≥ 7, « Confondu » plusieurs catégories ≥ 7, « non-identifié » = une ou plusieurs catégories ≤ 6.

5.5. CONCLUSIONS

Nous venons de présenter dans ce chapitre les résultats en perception des voyelles obtenus à partir de deux expériences : l'expérience de recherche des **prototypes** vocaliques et l'expérience de Catégorisation de l'espace vocalique.

En ce qui concerne la première expérience, l'utilisation de la **dynamique** a facilité la tâche des auditeurs des trois langues pour identifier les **prototypes** vocaliques de leur

système. Les espaces vocaliques et la dispersion intra-catégorie vocalique en fonction des trois contextes de présentation sont plus grands en **dynamique**, dans les trois langues. Systématiquement, les auditeurs des trois langues choisissaient les trois voyelles extrêmes /i a u/ dans des positions significativement plus extrêmes qu'en **statique**. C'est pourquoi nous obtenons un agrandissement de la taille des espaces vocaliques en **dynamiques**. Pour ce qui concerne la taille des ellipses pour chaque voyelle, l'utilisation de la **dynamique** a permis aux auditeurs des trois langues d'identifier les **prototypes** de leurs systèmes tout en prenant en compte les effets coarticulatoires des contextes de présentation. Les effets du contexte consonantique sur les stimuli **dynamiques** sont conformes à ceux obtenus en production des voyelles pour chaque système. Les résultats de la classification correcte de l'analyse discriminante permettent d'évaluer le rôle des indices utilisés dans la caractérisation des **prototypes** vocaliques par les auditeurs de chaque système. En AJ et FR, le taux moyen de classification correcte a été augmenté de 15% et de 2%, respectivement. En AM, il a baissé, en moyenne, de 8% en **dynamique**. L'augmentation observée en AJ et en FR est conforme à nos attentes. En effet, nous nous attendions à ce que l'utilisation de la **dynamique** « améliore » la précision des choix des **prototypes** des auditeurs des trois langues. En AM, en revanche, l'utilisation de la **dynamique** a baissée les taux de classification, car à notre sens, l'espace vocalique plus « centralisé » obtenu en AM et le rapprochement entre toutes les voyelles **prototypiques** et plus particulièrement les brèves, ont affecté la discrimination entre les voyelles de l'AM et par conséquent l'analyse discriminante n'a pas pu les séparer avec des taux supérieurs à ceux obtenus en **statique**. Pour la comparaison inter-langues, la **dynamique** a permis de différencier les trois langues avec un taux moyen de classification supérieur de 7%, conformément à nos attentes.

Concernant les résultats de la seconde expérience : pour la catégorisation de l'espace vocalique, les auditeurs des trois langues ont pu identifier les voyelles de leurs systèmes majoritairement. Les auditeurs marocains avaient des problèmes à catégoriser les voyelles brèves de leur système : l'utilisation de différents « représentants » d'une seule voyelle centrale les a perturbé et a conduit à des confusions importantes. Les auditeurs jordaniens ont été perturbés dans leurs catégorisations. La plupart du temps, ces confusions concernaient

trois zones : la zone des voyelles /i: i e:/, la zone /a a:/ et la zone /o: u u:/ . Nous avons observé que l'utilisation des stimuli avec des durées variables (i.e., longs à 200 ms et brefs à 100 ms) n'a pas aidé les auditeurs marocains ni jordaniens dans la catégorisation des voyelles brèves. À l'inverse, les auditeurs marocains ont catégorisé majoritairement certains stimuli brefs comme représentant la voyelle /i:/. Enfin, les auditeurs français étaient les seuls à pouvoir catégoriser leur espace vocalique sans confusions importantes (sauf pour les deux voyelles /ɑ ɔ/). Ceci indique que les auditeurs marocains et jordaniens semblent ne pas être « conscients » de leurs systèmes vocaliques brefs, car l'essentiel des confusions a été causé par les voyelles brèves dans les deux dialectes arabes. Nous pensons ainsi que les différences de statut relatif et du rôle des voyelles en arabe et en français sont à l'origine des difficultés rencontrées par les auditeurs arabophones. La durée vocalique semble ne pas être un indice primordial pour différencier entre les voyelles longues et brèves dans les deux dialectes arabes et plus spécifiquement pour les voyelles brèves. Pour les auditeurs marocains et jordaniens, ce sont les différences de qualité des timbres vocaliques qui sont à l'origine de la discrimination entre les voyelles de leurs systèmes.

Nous observons une amélioration de la précision de la catégorisation des espaces vocaliques dans les trois langues en utilisant les stimuli **dynamiques**. Les auditeurs des trois langues nous ont rapporté la facilité qu'ils avaient à catégoriser les espaces vocaliques de leurs systèmes avec les stimuli **dynamiques**, car, d'après eux, ces stimuli se rapprochaient davantage de la réalisation des voyelles catégorisées.

Les auditeurs marocains et jordaniens ont trouvé cette tâche difficile, surtout pour l'identification des voyelles brèves de leurs systèmes : /ɪ ɐ ʊ/ en AM et /i a u/ en AJ. En AM, les confusions de catégorisation de ces trois voyelles concernent le choix d'une des trois catégories comme « représentant » de la voyelle centrale /ə/. En AJ, en revanche, les confusions concernent la différenciation de ces trois voyelles de leurs correspondantes longues, où on a vu émerger trois zones différentes dans les catégorisations des auditeurs jordaniens : une zone /i: i e:/, une zone /a a:/ et une zone /o: u u:/. Les auditeurs français, quant à eux, n'ont pas pu différencier les deux voyelles /ɔ ɑ/ dans leurs catégorisations avec les deux modalités, en les confondant soit entre elles, soit avec /o a/, respectivement. Ceci

peut s'expliquer par le fait que, en fonction des résultats de production que nous avons obtenus, la voyelle /ɑ/ n'est plus réalisée en tant que telle en FR, mais confondue avec la voyelle /a/. Toutefois, la confusion dans la catégorisation de la voyelle /ɑ/ permet de voir qu'en perception, les auditeurs français perçoivent toujours la présence de la distinction /a α ɔ/.

CONCLUSION GÉNÉRALE ET PERSPECTIVES

Notre thèse avait pour double objectif de comprendre le rôle qu'ont les indices **dynamiques** (et plus particulièrement les différentes informations spectrales contenues dans les transitions formantiques) dans l'identification des voyelles en arabe et d'évaluer l'effet de la densité des systèmes sur la taille et la distribution des timbres dans les espaces vocaliques lors d'une comparaison avec le français.

Dans cette thèse, nous avons cherché à identifier la meilleure description de la voyelle en arabe qui permettait de comprendre le rôle de la caractérisation **dynamique** dans la perception des voyelles par les arabophones. Comme nous l'avons vu dans le premier chapitre, la voyelle en arabe a un statut particulier : les grammairiens arabes l'ont décrite comme une partie inséparable de la consonne. Elle ne peut exister, à l'intérieur ou à l'initial d'un mot, que si elle est associée à une consonne ou à un coup de glotte. La structure morphologique de l'arabe fait que les consonnes sont les seules porteuses du sens lexical, tandis que les voyelles portent le sens grammatical (voir Chapitre 1). Nous avons été amené à penser qu'une simple représentation **statique** de la voyelle en arabe (i.e., en suivant les propositions du modèle « *Simple Target* ») n'est peut-être pas suffisante. Les propositions de la littérature en ce qui concerne une représentation plus **dynamique** des voyelles nous a alors conduit à penser que l'analyse des effets coarticulatoires des consonnes adjacentes sur les voyelles (i.e., en suivant les propositions du modèle « *Dynamic Specification* ») peut contribuer à une meilleure appréhension du statut de la voyelle en arabe et de sa perception (voir Chapitre 2). Nous avons voulu également évaluer l'incidence de la **dynamique** dans la description acoustico-perceptive des voyelles en arabe et en français. En particulier, nous voulions savoir si les deux populations utilisent les mêmes indices.

Dans notre thèse, nous avons décidé de comparer l'arabe et le français, d'une part, à cause des différences morpho-phonologiques entre les deux langues (i.e., une morpho-phonologie différente pouvant affecter le statut des voyelles dans les deux langues), et d'autre part, afin d'évaluer le rôle de la densité des systèmes sur l'organisation des voyelles à l'intérieur de l'espace vocalique. Rappelons que, dans ce travail, nous avons étudié les

systèmes vocaliques de l'arabe marocain (avec un système à cinq voyelles /i: a: ə u u:/ au maximum), de l'arabe jordanien (avec un système à huit voyelles /i: i e: a: a o: u u:/) et du français (avec un système à onze voyelles orales /i e ε a α ɔ o u y ø œ/), suivant l'hypothèse que le nombre de voyelles dans une langue affectent la taille et l'organisation de l'espace vocalique. Trois visions différentes s'offrent à nous pour expliquer les effets de la densité des systèmes : i) la taille et l'organisation des voyelles à l'intérieur de l'espace vocalique sont invariant (suivant la **théorie quantique (TQ)**, Stevens, 1972, 1989, 2003), ii) l'espace vocalique est variable en termes de taille, d'organisation et de dispersion intra-catégorie (suivant la **théorie de la dispersion adaptée (TDA)**, Liljencrants & Lindblom, 1972 ; Lindblom, 1986) et enfin iii) l'espace vocalique est variable et focal (en termes de saillance spectrale entre les formants), (suivant la **théorie de la dispersion~focalisation (TDF)**, Schwartz *et al.*, 1989 ; Vallée, 1994 ; Boë *et al.*, 1996 ; Schwartz *et al.*, 1997_a, 1997_b), (voir Chapitre 3).

Afin de tester nos hypothèses, d'une part, sur le rôle des indices **dynamiques** potentiels en perception des voyelles, et de l'autre, pour évaluer les effets de la densité des systèmes sur l'organisation des voyelles dans les trois langues étudiées, nous avons cherché tout d'abord à identifier les différents paramètres **statiques** et **dynamiques** à partir des données de production des voyelles en arabe marocain, en arabe jordanien et en français. Une fois la pertinence des différents paramètres **dynamiques** évaluée, par rapport aux **statiques**, nous avons procédé à des tests de perception permettant de mettre en évidence le rôle des différents indices exploités.

Les résultats obtenus en production des voyelles ont été présentés dans le Chapitre 4. Nous avons proposé une description approfondie des systèmes vocaliques des trois langues étudiées. Nous avons présenté tout d'abord les résultats obtenus en ce qui concerne les effets de la densité des systèmes sur l'organisation des espaces vocaliques entre les trois langues, en évaluant d'une part, les différences entre les trois voyelles extrêmes /i a u/ (voyelles longues en arabe) dans les trois langues et de l'autre, les effets de la densité des systèmes sur l'organisation de la totalité des voyelles.

À partir des paramètres **statiques**, nos résultats ont montré que la taille des espaces vocaliques obtenue à partir des trois voyelles /i a u/ est significativement plus grande en français par rapport à celui obtenu en arabe marocain ou en arabe jordanien, lesquels ne présentent pas d'espaces significativement différents. Lorsque la comparaison des espaces est effectuée à partir de la totalité des voyelles dans chaque langue, nous avons trouvé que la taille de l'espace vocalique en français est significativement plus grande que celle en arabe jordanien qui à son tour présente un espace vocalique significativement plus grand que celui obtenu en arabe marocain. Ces différences vont dans le sens de l'hypothèse d'une expansion des espaces vocaliques en fonction de la densité des systèmes, et confortent ainsi les propositions de la **TDA**. Lorsque nous avons comparé la taille des ellipses de dispersion pour chaque voyelle, nos résultats ont montré que c'est l'arabe marocain qui présente les ellipses les plus petites, tandis que c'est l'arabe jordanien qui obtient les ellipses les plus grandes. Ce résultat va dans le sens des propositions de la **TDA**, car nous obtenons deux stratégies différentes : nous obtenons une réduction de la taille de l'espace vocalique en arabe marocain, tandis que la dispersion intra-catégorie vocalique reste faible afin de garder un maximum de distinctivité entre les voyelles de l'arabe marocain.

Nous avons demandé aux locuteurs des trois langues de produire les voyelles de leurs systèmes dans des contextes Mot et Syllabe et enfin en Isolation, en essayant de conserver la qualité de production de la voyelle en Mot (e.g., [bo:se~bo:~o:] = "bisou" en AJ). L'utilisation de ces trois situations de production permettaient d'obtenir trois réalisations différentes de la voyelle afin d'évaluer les différents degrés de réduction vocalique en fonction du contexte environnant (i.e., réduction vocalique maximale en situation Mot vs réduction vocalique minimale – voire nulle – dans la situation Isolation). Ces différents environnements nous ont permis d'évaluer ces différentes situations selon l'explication donnée Björn Lindblom (Lindblom, 1990) : en partant de la forme *Hypo-Articulation* à l'*Hyper-Articulation* : la production des voyelles en Mot correspond relativement à la forme *Hypo-Articulée* où une variabilité intra- et inter-individuelle importante est présente ; tandis que la forme *Hyper-Articulée* naît de la production des voyelles en Isolation. Les résultats obtenus en fonction de la situation de production montrent que pour l'arabe marocain et le français les espaces

vocaliques obtenus en Isolation sont significativement plus grands que ceux en Mot et ce quel que soit le lieu d'articulation ; ceux en Syllabe se situent entre les deux. En arabe jordanien, en revanche, aucune différence significative n'a été obtenue en fonction des situations de production. Les résultats obtenus en arabe marocain et en français vont ainsi dans le sens des propositions de Lindblom (1990) sur l'effet de la quantité d'information contextuelle contenue dans le signal acoustique sur la réalisation plus ou moins invariante des voyelles ; les résultats en arabe jordanien ne confirment pas cette hypothèse. Ce résultat en arabe jordanien peut être expliqué par les difficultés rencontrées par les locuteurs lors de la production des voyelles dans les situations Syllabe et Isolation : certains locuteurs jordaniens produisaient les voyelles fermées et mi-fermées de leur système comme une voyelle [ə] ou [a], car d'après eux, cette tâche était difficile à réaliser.

En ce qui concerne les résultats obtenus avec les paramètres **dynamiques**, nous avons émis l'hypothèse d'un effet de la densité des systèmes sur la valeur des **pentés formantiques**. Nous avons supposé que plus une langue a de voyelles, plus la réalisation de celles-ci correspondrait à une forme *Hyper-Articulée* et plus la **pen**te de ses transitions serait importante. Nos résultats ont montré que les **pentés formantiques** obtenues en français sont significativement plus « raides » que celles en arabe marocain ; l'arabe jordanien se situe plutôt entre les deux. Les différences inter-langues observées ne se limitent pas à la **pen**te. Nos résultats ont montré également que, pour une même voyelle produite dans une même situation de production, le point de départ des transitions (ou l'**onset** vocalique dérivé), la **pen**te des transitions et la « cible vocalique » dérivée pour un formant précis, sont significativement différents entre les trois langues. Autrement dit, dès le relâchement de l'occlusion, on dispose d'indices nécessaires pour identifier la voyelle, ainsi que son environnement consonantique, la situation de production et la langue produisant cette voyelle a été produite. Nos résultats ont montré également que les **pentés formantiques** obtenues en fonction des trois situations de production sont significativement différentes : en Mot, la **pen**te formantique est la plus « raide » et en Isolation, est la plus « aplatie » quels que soient le formant étudié et le lieu d'articulation. De plus, nous avons montré qu'en Isolation, des **pentés formantiques** non-nulles sont obtenues, confirmant ainsi l'hypothèse selon laquelle

les voyelles produites en Isolation présentent des trajectoires **dynamiques intrinsèques**. En ce qui concerne les résultats des **pentés formantiques** obtenues en fonction des lieux d'articulation, nos résultats ont montré leur conformité avec les propositions de la théorie du « locus virtuel » (Couper *et al.*, 1952 ; Delattre *et al.*, 1955) : toutes les transitions formantiques trouvent leurs origines à un point virtuel situé soit autour de 1000 Hz, pour le lieu bilabial, soit autour de 1800 Hz, pour le lieu dental, soit autour de 1300 Hz s'il s'agit d'une voyelle postérieure ou autour de 3000 Hz s'il s'agit d'une voyelle antérieure, pour le lieu vélaire. À partir des **pentés formantiques** (obtenues par des analyses de régressions linéaires et polynomiales de 2^{ème} et de 3^{ème} ordre), on est également capable de modéliser le point de départ, la **penne** et la « cible vocalique » dérivée, pour n'importe quelle voyelle et ce quels que soient le formant étudié, le lieu d'articulation, la situation de production, le locuteur et la langue. Ces résultats sont homogènes avec ou sans normalisation de durée. En effet, nous avons proposé de modéliser les **pentés formantiques** dans les trois langues en normalisant les durées des transitions, qui sont variables entre les voyelles longues de l'arabe et les « brèves » du français. La littérature explique que la durée vocalique peut conditionner le degré de réduction vocalique (Lindblom, 1963_a) et par conséquent le degré d'inclinaison des **pentés** (voir en revanche, Gay, 1978 ; Pols & van Son, 1993, pour des effets inverses). Dans notre cas, nous nous attendions à observer son influence de la manière suivante : plus la durée vocalique est longue, mieux la cible vocalique est atteinte et moins l'inclinaison de la transition formantique est forte. À durée variable, nous devrions nous attendre alors à observer des valeurs de **pentés formantiques** plus élevées en FR et plus basses en AM ou en AJ (i.e., degré d'inclinaison plus forte en FR et plus faible en AM ou en AJ). Nous avons décidé alors de normaliser les durées pour étudier simplement la valeur de chaque **penne** indépendamment de la longueur de la transition. Les résultats des **pentés formantiques** normalisées montrent que le français ne présente pas de différences par rapport à ceux obtenus en arabe marocain ou en arabe jordanien, allant ainsi dans le sens des propositions de Lindblom (1963_a). Toutefois, nous avons montré que les points de départ des transitions et les « cibles vocaliques » dérivées des trois voyelles /i a u/, sont significativement différents entre les trois langues, indiquant ainsi que la densité des systèmes semble les affecter.

L'évaluation des résultats obtenus à partir des paramètres **statiques** et **dynamiques** a été conduite par une analyse discriminante, qui avait pour but de proposer une classification des voyelles dans les trois langues, à partir d'un paramètre précis (comme les fréquences formantiques, les durées vocaliques, etc...). L'utilisation de cette méthode de classification nous permet, d'une part, de vérifier la validité de nos résultats, et d'autre part, de savoir quelle est la meilleure façon de décrire les voyelles dans les trois langues à partir des différents paramètres utilisés. Les taux moyens de classification correcte montrent que le paramètre de la durée vocalique semble être le facteur primordial pour discriminer les trois langues. En effet, dès que ce paramètre est ajouté aux fréquences formantiques, i.e., les paramètres **statiques**, les taux de classification correcte augmentent significativement de 10 à 30 % (selon les cas). Ceci indique que la durée intrinsèquement différente entre les voyelles /i a u/ dans les trois langues (voyelles longues en arabe) a permis de séparer leurs voyelles. Cet effet est normal, car les voyelles longues en arabe présentent parfois des durées vocaliques deux fois supérieures à celles en français. Les **pentés formantiques** normalisées ont permis néanmoins de discriminer les trois langues avec des taux relativement élevés.

L'utilisation des différentes analyses de régression avait pour but de vérifier laquelle de ces méthodes est la mieux adaptée pour modéliser les trajectoires formantiques **intrinsèques** et **extrinsèques**. Nos résultats ont montré que l'utilisation de la régression linéaire suffisait à obtenir une modélisation « correcte » de ces trajectoires pour les formants F_1 et F_3 . Pour ce qui est du formant F_2 , nos résultats ont montré que l'analyse de régression polynomiale de 2^{ème} ordre propose une modélisation plus fidèle de ses trajectoires. Les taux de classification correcte ont ainsi augmenté significativement en utilisant les coefficients de la régression polynomiale de 2^{ème} ordre, indiquant ainsi que la modélisation des trajectoires formantiques obtenue à partir de cette analyse permet de mieux différencier les voyelles dans les trois langues. Les taux de classification des trois voyelles /i a u/ produites dans l'environnement /d^ʕ/ en arabe marocain et en arabe jordanien, sont très élevés et plus particulièrement lorsque les analyses **dynamiques** sont utilisées (avec et sans normalisation de durée). Ceci indique que des différences de production de ces voyelles dans l'environnement /d^ʕ/ existent entre les différents dialectes arabes et permettent de les

discriminer. La comparaison des taux de classification obtenus dans les trois langues à partir des paramètres **statiques** et **dynamiques** montre que ce sont ces derniers qui permettent le mieux de discriminer les voyelles dans les trois langues. L'augmentation de l'ordre de l'analyse de régression polynomiale du 2^{ème} au 3^{ème} semble par contre ne pas améliorer les taux de classification et rejoint ainsi les observations de Kirsty McDougall (McDougall, 2006), où la modélisation des trajectoires formantiques des monophthongues semble non pertinente en utilisant la régression polynomiale de 3^{ème} ordre (qui est cependant plus adaptée à la modélisation des diphtongues).

En ce qui concerne l'étude des spécificités de chaque langue, nos résultats ont montré des concordances et des désaccords avec les propositions de la littérature. En ce qui concerne les systèmes vocaliques de l'arabe marocain, les deux voyelles brèves /ə u/ sont produites différemment en fonction des environnements consonantiques : elles semblent fusionner dans les environnements /d d^ʕ/ ; en revanche elles sont réalisées comme deux voyelles distinctes dans les environnements /b k/. Les **pentés formantiques**, les points de départ des transitions et les « cibles vocaliques » dérivées obtenues pour ces deux voyelles sont significativement différentes quels que soient le lieu d'articulation et la situation de production. Les résultats de l'analyse discriminante montrent la possibilité de discriminer ces deux voyelles avec des taux dépassant 60% (et au-delà de 90% dans l'environnement /k/), à partir des différents paramètres utilisés. Ces résultats montrent que lors de leurs productions, les locuteurs marocains produisent deux voyelles significativement différentes sur le plan **dynamique** ; en **statique**, elles sont significativement différentes dans les environnements /b k/. Le système vocalique que nous proposons en arabe marocain, à partir de nos données de production, correspond à un système à cinq voyelles /i: a: ə u u:/.

La littérature propose un système vocalique pour l'arabe jordanien de huit voyelles : /i: i e: a a: o: u u:/. Les résultats obtenus à partir des paramètres **statiques** montrent que les deux voyelles brèves /i u/ sont produites différemment de leurs correspondantes longues, et sont positionnées respectivement aux côtés des voyelles longues /e: o:/, et ce quels que soient le lieu d'articulation et la situation de production (sauf dans l'environnement /d/). La voyelle brève /a/ est produite phonétiquement comme une voyelle mi-ouverte [ɐ] seulement dans la

situation de production Mot ; dans les situations Syllabe et Isolation, elle obtient un timbre équivalent à une voyelle ouverte [a]. En ce qui concerne les paramètres **dynamiques**, nos résultats ont montré que les deux voyelles brèves /i u/ sont différentes de leurs correspondantes longues en fonction : i) des points de départ des transitions, ii) des **pentés** des transitions et iii) des « cibles vocaliques » dérivées. Pour la voyelle /a/, le même cas de figure que celui en **statique** est obtenu. Les résultats de l'analyse discriminante ont confirmé nos résultats, car les deux voyelles brèves /i u/ ont été la plupart du temps confondues avec les deux voyelles /e: o:/, respectivement. Ces résultats nous permettent de proposer le système vocalique en arabe jordanien comme ayant la structure suivante : /i: e e: a a: o: o u/.

Les description des systèmes vocaliques du français propose onze voyelles orales /i e ε a α ɔ o u y ø œ/. Les résultats que nous avons obtenus ont montré que les locuteurs français produisent les deux voyelles ouvertes /a α/ de la même manière (i.e., aucune différence significative n'a été obtenue), en utilisant les paramètres **statiques** et **dynamiques**. En ce qui concerne les voyelles dites « médianes » /e ε o ɔ ø œ/, nos résultats ont montré qu'il est possible de les différencier les locuteurs français, en **statique** et en **dynamique**. La classification des voyelles en français a permis de séparer ces voyelles, mais les confusions sont très importantes plus particulièrement pour les paires de voyelles /ø œ/ et /o ɔ/.

La classification des voyelles pour chacune des trois langues a permis de vérifier l'apport de la **dynamique** dans la caractérisation des trajectoires formantiques par rapport aux paramètres **statiques**. En arabe marocain, l'ajout de la durée vocalique aux paramètres **statiques** a amélioré significativement les taux de classification dans les deux situations de production Mot et Syllabe, en revanche, en situation Isolation, la modélisation des trajectoires **intrinsèques** avec la régression polynomiale de 3^{ème} ordre a amélioré significativement les taux de classification (BinomialQ ; p<0,05). En arabe jordanien, la durée vocalique ajoutée aux paramètres **statiques**, et les coefficients de la régression linéaire, semblent être les deux paramètres les plus pertinents pour différencier les voyelles dans les situations de production Mot et Syllabe. En situation Isolation, les trois analyses de régression : linéaire, polynomiales de 2^{ème} et de 3^{ème} ordre modélisent de la même façon les trajectoires **intrinsèques** des formants vocaliques, par conséquent nous supposons qu'une simple analyse de régression

linéaire est suffisante pour la modélisation de ces trajectoires **intrinsèques**. En français enfin, et dans les deux situations Mot et Syllabe, l'utilisation de la régression linéaire donne une bonne caractérisation **dynamique** des trajectoires formantiques **extrinsèques**. En Isolation, ce sont les deux régression linéaire et polynomiale de 2^{ème} ordre qui modélisent le mieux les trajectoires formantiques **intrinsèques**.

Les résultats obtenus dans les trois langues montrent que les différents paramètres **dynamiques** apportent une bonne caractérisation des trajectoires formantiques **intrinsèques** et **extrinsèques** puisqu'ils ont permis de discriminer les voyelles de chaque système avec des taux significativement supérieurs à ceux obtenus en utilisant les paramètres **statiques**. Notre hypothèse de départ expliquait que c'est en arabe principalement que la caractérisation **dynamique** des voyelles pouvait apporter plus de précision lors de leurs descriptions (évaluée à partir des analyses discriminantes), or nos résultats ont montré que l'amélioration est observée dans les trois langues.

La deuxième partie de notre travail consistait à évaluer le rôle des différents indices **dynamiques** potentiels identifiés à partir des données de production dans la perception des voyelles dans les trois langues. Nous avons ainsi proposé deux expériences de perception des voyelles aux auditeurs des trois langues : une expérience de recherche des **prototypes** vocaliques dans chaque langue (à partir d'une *Méthode d'Ajustement de Formants*, MOA, Johnson *et al.*, 1993), et une expérience de catégorisation des espaces vocaliques.

Pour ce qui est des **prototypes** vocaliques identifiés en arabe marocain, les résultats montrent que l'utilisation de la **dynamique** a facilité la tâche des auditeurs marocains lors de l'identification des **prototypes** de leur système. En ce qui concerne les voyelles brèves de l'arabe marocain, nous avons décidé de tester le statut phonologique de ces anciennes voyelles héritées de l'arabe classique dans la variété actuelle de l'AM, i.e., /i a u/. La littérature explique que les deux voyelles */i/ et */a/ ont fusionné et forment désormais une seule voyelle centrale /ə/, tandis que la voyelle */u/ est resté en opposition quantitative avec sa correspondante longue. Dans nos résultats de production, nous avons obtenu des ellipses de dispersion pour la voyelle /ə/ allant de la voyelle [ɪ] à la voyelle [ɐ]. Cette variation de qualité

de voyelle est affectée par les différences de comportement des locuteurs marocains. Nous avons pensé alors que, puisque certains locuteurs produisent la voyelle /ə/ avec des variations allophoniques importantes, son statut phonologique n'est peut-être pas bien défini. Nous avons ainsi proposé de tester *perceptivement* le statut des trois voyelles brèves /ɪ ʊ ʊ/. Les résultats obtenus ont montré que les auditeurs marocains sont sensibles à cette différence, car ils choisissent deux voyelles prototypiques différentes pour la seule catégorie /ə/ et une seule voyelle prototypique pour la catégorie /ʊ/, à partir des paramètres **statiques** et **dynamiques**. Nous avons supposé que la structure syllabique utilisée, i.e., CVC, pour la plupart des items testés, pouvait affecter la qualité du timbre de ces voyelles. En effet, dans cette structure de type CVC, les voyelles brèves peuvent être réalisées comme des voyelles « non-réduites », car l'accent tombe sur cette syllabe, et par conséquent sa perception peut être affectée. Toutefois, nous pensons que l'effet de l'accent et de la structure syllabique CVC peut affecter seulement les voyelles **prototypiques** identifiées à partir des stimuli **dynamiques**, car toute l'information coarticulatoire nécessaire à leurs identifications est présente dans les stimuli. En revanche, avec les stimuli **statiques**, aucune information n'est disponible et par conséquent les effets d'accentuation peuvent être moins visibles. Les résultats de perception avec les stimuli **statiques** et **dynamiques** ont montré les mêmes effets, i.e., trois voyelles **prototypiques** brèves ont été identifiées. Nous pouvons supposer que la « représentation mentale » de ces trois voyelles brèves est différente pour permettre l'identification de trois prototypes différents.

Les résultats obtenus à partir de la seconde expérience de perception (catégorisation des espaces vocaliques), ont montré que les auditeurs marocains sont capables de catégoriser correctement les voyelles longues de leur système ; les trois voyelles brèves n'ont, par contre, pas été catégorisées. Les auditeurs marocains nous ont expliqué que, systématiquement, ils ne savaient pas quoi choisir comme catégorie pour les trois voyelles brèves, car ils les confondaient tout le temps. Il semble donc que nous avons proposé aux auditeurs marocains trois « représentants » différents d'une « possible même catégorie vocalique », i.e., /ə/. Les variations entre les trois « représentants » obtenus lors de la première expérience peuvent être corrélées au fait que les effets coarticulatoires des consonnes ont été bien assimilés par les

auditeurs marocains et leurs choix en étaient affectés. D'autres tests perceptifs devraient être envisagés pour évaluer le statut phonologique des voyelles de l'arabe marocain et plus particulièrement les voyelles brèves.

Les résultats obtenus en arabe jordanien à partir de la première expérience de perception sont conformes aux données de production, car les **prototypes** vocaliques identifiés pour les deux voyelles brèves /i u/ sont significativement différents de ceux identifiés pour leurs correspondantes longues. Ces **prototypes** se rapprochent respectivement des deux voyelles longues /e: o:/. Les résultats de l'analyse discriminante ont montré que les deux voyelles brèves /i u/ sont toujours confondues avec les deux voyelles longues /e: o:/, respectivement et rarement avec leurs correspondantes longues. Ces résultats sont homogènes quel que soit le paramètre utilisé. Nous avons remarqué néanmoins que l'utilisation de la **dynamique** a énormément facilité la tâche des auditeurs jordaniens, car nous obtenons plus de précision dans les choix et, de plus, la taille des aires de dispersion des voyelles est significativement plus petite en **dynamique**. L'utilisation des stimuli **statiques** a, semble-t-il, perturbé à nouveau les auditeurs jordaniens, car une grande variabilité de réponses intra- et inter-sujets est apparue. Les taux de classification correcte de l'analyse discriminante ont été améliorés avec la **dynamique**, indiquant que les voyelles choisies par les auditeurs jordaniens sont significativement différentes et statistiquement contrastées. En ce qui concerne les résultats obtenus à partir de la seconde expérience de perception, les auditeurs jordaniens ont été souvent très perturbés dans leur catégorisation. Ainsi, nous avons pu dégager trois zones de confusion en arabe jordanien : la zone des voyelles /i: i e:/, la zone des voyelles /a a:/ et la zone des voyelles /o: u u:/. Les voyelles brèves n'ont jamais été correctement catégorisées, les longues quant à elles, ont été quelquefois bien catégorisées. Nous nous posons ainsi des questions sur le statut phonologique des trois voyelles brèves en arabe jordanien : constituent-elles trois voyelles à part entière ou les « représentations mentales » les impliquant sont-elles amalgamées avec celles de leurs correspondantes longues ?

Enfin, les résultats obtenus en français, montrent que les **prototypes** vocaliques choisis à partir des stimuli **statiques** et **dynamiques** ne sont pas significativement différents. En effet, nous n'avons pas pu mesurer l'effet de la **dynamique** sur les réponses des auditeurs

français, car la taille des aires de dispersion et des ellipses de chaque voyelle, obtenue dans les deux modalités est comparable. Les résultats de l'analyse discriminante ont montré une légère amélioration du taux de classification correcte en **dynamique**. Les auditeurs français nous ont pourtant expliqué que la tâche d'identification des voyelles était plus simple dans l'expérience avec des stimuli **dynamiques**. En ce qui concerne les voyelles « problématiques », les auditeurs français ont souvent choisi des **prototypes** non différents significativement pour les deux voyelles ouvertes /a a/. Pour ce qui est des voyelles « médianes », les **prototypes** vocaliques identifiés par les auditeurs français sont significativement différents. Ces résultats vont dans le sens de ceux obtenus en production. En ce qui concerne les résultats de la seconde expérience, les auditeurs français sont les seuls parmi les trois populations à pouvoir catégoriser correctement leur espace vocalique. Toutes les voyelles, ont été très bien catégorisées, sauf les deux voyelles /ɔ a/ qui ont été souvent confondues soit entre elles soit avec les voyelles /o a/, respectivement.

À partir de l'expérience d'ajustement de formants, nous avons effectué une comparaison inter-langues en évaluant les choix **prototypiques** des trois voyelles extrêmes /i a u/ (voyelles longues en arabe). La littérature explique que, puisque la perception de la parole est un phénomène *many-to-one* où une multitude de stimuli acoustiques sont identifiés comme un seul stimulus perceptif, l'identification des voyelles dans les langues ne diffèrera pas : les auditeurs choisissent des voyelles **prototypiques** non différentes (Meunier, 2005). D'après les propositions de « *The Native Language Magnet* » (Kuhl & Iverson, 1995*), l'espace acoustico-perceptif est divisé en plusieurs catégories à partir de frontières psycho-acoustiques qui sont fonction du système phonologique de chaque langue. La taille de l'espace ne change pas, en revanche l'organisation des voyelles à l'intérieur de l'espace vocalique peut être différente. Les résultats que nous avons obtenus ont montré que la taille de l'espace acoustico-perceptif est fonction de la densité des systèmes. Nous obtenons un espace perceptif significativement plus petit en arabe marocain et plus grand en français, en **statique** et en **dynamique**. Aucune différence significative n'est obtenue entre l'arabe marocain et l'arabe jordanien. L'utilisation de la totalité des voyelles donne la même tendance, mais la différence entre l'arabe marocain et l'arabe jordanien devient significative. Ainsi, c'est l'arabe

marocain qui obtient l'espace le plus petit, le français le plus grand, et l'arabe jordanien entre les deux.

Globalement, les résultats obtenus en production et en perception sont homogènes en ce qui concerne les effets de la **dynamique**. Nous avons montré qu'en plus des dialectes arabes, la caractérisation **dynamique** des voyelles produites en français est pertinente pour leur description. Les auditeurs marocain et jordanien sont sensibles aux variations apportées par les stimuli **dynamiques** et plus particulièrement pour ce qui est des effets coarticulatoires des consonnes ; les français nous ont expliqué la facilité rencontrée avec la **dynamique** : facilité pourtant non validée statistiquement. Ces résultats montrent que la prise en compte des aspects **dynamiques** en production et en perception des voyelles est un facteur primordial pour la recherche des unités de base en production et en perception de la parole. Ces résultats sont en accord avec les propositions des nouvelles théories en perception de la parole et plus particulièrement avec le modèle à exemplaires (pour un résumé des modèles à exemplaires, voir Nguyen, 2005 ; Nguyen *et al.*, 2007). Les partisans des modèles à exemplaires expliquent que l'auditeur a la capacité de stocker en mémoire les différentes formes de surface associées à un mot. Les variations phonétiques sont alors décodées directement par l'auditeur afin que le meilleur candidat pour une catégorie – vocalique – soit choisi. En nous inspirant de ces modèles, nous pensons en effet que les auditeurs peuvent utiliser les paramètres **dynamiques** obtenus à partir de la modélisation des trajectoires **extrinsèques** et ainsi retrouver le « représentant mental » de la catégorie examinée. En **statique**, les seules trajectoires **intrinsèques** ne sont peut-être pas suffisantes pour identifier la « représentation mentale » d'une voyelle, car les chevauchements sont plus importants.

Nous pouvons supposer également que les paramètres **statiques** et **dynamiques** disponibles dans le signal et adaptés selon la langue, le contexte, etc... soient pris en compte par les auditeurs. La vision apportée par René Carré (Carré, 2006, 2007 (à paraître) ; Carré *et al.*, 2004), mais également par Winifred Strange (Strange, 1989, 1999_a) va dans le sens de cette prise en compte des informations **statiques** et **dynamiques** et plus spécifiquement sur les aspects **intrinsèques** et **extrinsèques** des trajectoires formantiques. Nous avons ainsi voulu savoir quel est l'apport des informations fournies par les deux analyses **statiques** et

dynamiques dans la séparation des voyelles /i a u/ dans les trois langues. Les taux de classification obtenus n'étaient pas significativement plus élevés lorsque les paramètres **statiques** et **dynamiques** ont été utilisés en comparaison avec la **dynamique** seule. Ceci indique que l'information apportée par les deux indices est complémentaire.

Vu que les auditeurs sont capables d'identifier avec « exactitude » les « représentants » de chaque voyelle en utilisant la **dynamique**, pourquoi les auditeurs arabophones ont-ils exprimé des difficultés lors de la tâche de catégorisation des espaces vocaliques ; difficultés absentes lors de l'identification des **prototypes** ? En effet, lors de la première tâche, les auditeurs arabophones, marocains et jordaniens, arrivaient facilement à identifier les **prototypes** vocaliques correspondants aux voyelles brèves de leurs systèmes, tandis qu'avec l'expérience de catégorisation, les confusions étaient multiples. Est-ce que les voyelles brèves identifiées avec exactitude par les auditeurs arabophones dans l'expérience d'identification des **prototypes** n'était qu'un artefact de la tâche : les auditeurs étaient « obligés » d'identifier un prototype pour chaque voyelle, tandis que pour la tâche de catégorisation, ils avaient le choix entre différents « représentants » et la case « aucun ne correspond » ? Peut-on supposer que la tâche d'identification des prototypes rend davantage compte des « variations phonétiques » existantes entre les stimuli proposés (i.e., un phénomène *one-to-many*) tandis que la tâche de catégorisation rend davantage compte de la « représentation » de ces stimuli pour chaque auditeur (i.e., un phénomène *many-to-one*) ? Si c'était le cas, pourquoi aucune difficulté n'est rencontrée par les auditeurs français ? Nous pensons en effet que le statut phonologique des voyelles en arabe peut être à l'origine des difficultés rencontrées par les auditeurs arabophones, surtout en catégorisation de l'espace. Il serait judicieux d'effectuer de nouveaux tests perceptifs : d'ajustement de formant et de catégorisation de l'espace vocalique, où une séparation de la catégorisation des voyelles longues et brèves dans deux tâches différentes permettrait de rendre compte du statut des voyelles. L'utilisation d'autres environnements consonantiques (en plus de l'environnement /d/ utilisé en catégorisation) peut contribuer à une meilleure compréhension des différences « représentationnelles » au niveau de ces voyelles.

Le travail que nous avons fourni dans cette thèse contribue à une meilleure connaissance de la voyelle et de son statut en arabe. Nous avons procédé à un travail de description acoustico-perceptive conséquent pour deux dialectes arabes, et si les corpus élaborés sont de taille conséquente, ils restent avant tout de la parole contrôlée. Il serait intéressant de comparer nos résultats à des résultats obtenus à partir de parole spontanée, où des chevauchements et des réductions vocaliques plus importantes sont disponibles. Une comparaison qui inclut la production des voyelles comme réalisées en Isolation, en Syllabe, en Mot, en Phrase et en parole spontanée, permettrait une meilleure appréhension de l'effet de la quantité d'information contextuelle sur le degré plus ou moins invariant de production des voyelles. Nous souhaitons également approfondir l'étude de la variabilité intra- et inter-individuelle qui permettrait d'identifier les différences de stratégies individuelles en production et en perception. L'étude du rapport existant entre les paramètres **statiques** et **dynamiques** que nous avons mis en évidence dans notre travail permet de mieux comprendre le rôle de chacun de ces paramètres en perception des voyelles. Ces deux types de paramètres peuvent être utilisés simultanément et d'une façon complémentaire par les auditeurs de différentes langues. Une étude plus poussée des stratégies individuelles rendra davantage compte de leurs rôles.

BIBLIOGRAPHIE

RÉFÉRENCES CONSULTÉES

- Abd-el-Jawad, H. R., and Awwad, M., 1989, "Reflexes of Classical Arabic Interdentals: A Study in Historical Sociolinguistics", in *Linguistische Berichte*, Vol. 122, pp. 259-282.
- Abou-Haidar, L., 1991, "*Variabilité et Invariance du Système Vocalique de l'Arabe Standard*", Thèse de Doctorat (NR), Université de Franche-Comté, Besançon, France.
- Al-Ani, S., 1970, "*Arabic Phonology: An Acoustical and Physiological Investigation*", The Hague Mouton, Paris.
- Alioua, A., 1987, "*Étude Phonétique et Morphologique des Voyelles et des Diphtongues de l'Arabe Littéral Moderne : analyse acoustique fondée sur la spectrographie (locuteurs marocains)*" Thèse de Doctorat, Université de Strasbourg II, Strasbourg, France.
- Al-Khatib, M. A. A., 1988, "*Sociolinguistic Change in an Expanding Urban Context: A Case Study of Irbid City, Jordan*", Ph.D. Dissertation, University of Durham, UK.
- Al-Sughayer, K. I., 1990, "*Aspects of Comparative Jordanian and Modern Standard Arabic Phonology*", Ph.D. Dissertation, Michigan State University, Michigan, USA.
- Al-Tamimi, J., 2001, "*Description linguistique de 6 parlers jordaniens : étude des interférences en production entre l'arabe dialectal jordaniens et l'arabe classique ; une approche phonético-phonologique et acoustique*", Mémoire de Maîtrise en Sciences du Langage, Université Lyon 2, Lyon, France.
- Al-Tamimi, J., 2002, "*Variabilité phonétique en production et en perception de la parole : le cas de l'arabe jordano-palestinien*", Mémoire du DEA en Sciences du Langage, Université Lyon 2, Lyon, France.
- Al-Tamimi, J., 2004, "L'équation du locus comme mesure de la coarticulation VC et CV : étude préliminaire en Arabe Dialectal Jordanien", in *Actes des XXV^{èmes} Journées d'études sur la Parole*, Fès, Maroc, pp. 9-12.
- Al-Tamimi, J., 2006, "Analyse dynamique de la réduction vocalique en contexte CV à partir des pentes formantiques en arabe dialectal et en français", in *Actes des XXVI^{èmes} Journées d'études sur la Parole*, Dinard, France, pp. 357-360.
- Al-Tamimi, J., 2007, "Static and dynamic cues in vowel production: A cross dialectal study in Jordanian and Moroccan Arabic", to appear in *Proc. of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, Germany.
- Al-Tamimi, J., and Barkat-Defradas, M., 2002, "Inter-dialectal and inter-individual variability in production and perception: A preliminary study in Jordanian and Moroccan Arabic", in *Proc. of the 5th International Conference of AIDA (Association Internationale de Dialectologie Arabe)*, Cadix, Espagne, pp. 171-186.
- Al-Tamimi, J., and Ferragne, E., 2005, "Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French", in *Proc. of InterSpeech*, Lisbon, Portugal, pp. 2465-2468.
- Al-Tamimi, J., Carré, R., and Marsico, E., 2004, "The status of vowels in Jordanian and Moroccan Arabic: Insights from production and perception", in *Journal of the Acoustical Society of America (JASA)*, Vol. 116:4, p. 2629A.
- Al-Tamimi, J., Girard, M., and Marsico, E., 2002, "Variabilité inter-langue et inter-individuelle en production et en perception : étude préliminaire en arabe dialectal et en français", in *Actes des XXIV^{èmes} Journées d'Étude sur la Parole*, Nancy, France, pp. 169-172.

- Al-Wer, E., 2000, "Rising of /a/ and related vocalic movements in the merging dialect of Amman", in *Proc. of the 3rd International Conference of AIDA (Association Internationale de Dialectologie Arabe)*, Malta, pp. 77-83.
- Al-Wer, E., 2002_a, "Jordanian and Palestinian dialects in contact: Vowel raising in Amman", in *Language Change: The Interplay of Internal, External, and Extra-Linguistic Factors (Contributions to the Sociology of Language, 86)*, M. C. Jones, and E. Esch (Eds), Mouton de Gruyter, Berlin ; New York, pp. 63-80.
- Al-Wer, E., 2002_b, "Education as a Speaker Variable" in *Language Contact and Language Conflict in Arabic. Variations on a Sociolinguistic Theme*, A. Rouchdy (Ed), Routledge, Curzon, London, UK, pp. 41-53.
- Amara, M. H., and Spolsky, B., 1996, "The construction of identity in a divided Palestinian village: Sociolinguistic evidence", in *Language and Identity in the Middle East and North Africa*, Y. Suleiman (Ed), Richmond Surrey, Curzon, pp. 81-99.
- Angoujard, J.-P., 1990, "*Metrical Structure of Arabic*", Foris, Dordrecht.
- Assmann, P. F., Nearey, T. M., and Hogan, J. T., 1982, "Vowel identification: Orthographic, perceptual, and acoustic aspects", in *Journal of the Acoustical Society of America (JASA)*, Vol. 71:4, pp. 975-989.
- Ayoub, G., and Bohas, G., 1983, "Les grammairiens arabes, la phrase nominale et le bon sens", in *The History of Linguistics in the Near East*, K. Versteegh, K. Koerner and H.-J. Niederehe (Eds), J. Benjamins, Amsterdam, Pays Bas, pp. 31-48.
- Bani-Yassin, R., and Owens, J., 1987, "The Phonology of a Northern Jordanian Arabic Dialect", in *Zeitschrift der Deutschen Morgenlandischen Gesellschaft*, Vol. 137:2, pp. 297-331.
- Barkat, M., 2000, "*Détermination d'indices acoustiques robustes pour l'identification automatique des parlers arabes*", Thèse de Doctorat en Sciences du Langage, Université Lyon 2, Lyon, France.
- Barkat-Defradas, M., Al-Tamimi, J., and Benkirane, T., 2003, "Phonetic variation in production and perception of speech: A comparative study of two Arabic dialects", in *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain, pp. 857-860.
- Belkaid, Y., 1984, "Les voyelles de l'arabe littéraire moderne : analyse spectrographique", in *Travaux de l'Institut de Phonétique de Strasbourg*, Vol. 16, pp. 217-240.
- Benkirane, T., 1981-82, "Durée, prosodie et syllabation en Arabe Marocain", in *Travaux de l'Institut de Phonétique d'Aix*, Vol. 8, pp. 49-83.
- Benkirane, T., 1982, "*Étude Phonétique et Fonctions de la Syllabe en Arabe Marocain*", Thèse de 3^{ème} Cycle de Phonétique Expérimentale, Fonctionnelle et Appliquée, Université de Provence, Aix-Marseille I, Aix-en-Provence, France.
- Benkirane, T., 2000, "*Codage Prosodique de l'énoncé en arabe marocain*", Thèse d'état, Université Aix-en-Provence, Aix-en-Provence, France.
- Blumstein, S., and Stevens, K., 1979, "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants", in *Journal of the Acoustical Society of America (JASA)*, Vol. 66:4, pp. 1001-1017.
- Blumstein, S., and Stevens, K., 1979, "Perceptual invariance and onset spectra for stop consonants in different vowel environments", in *Journal of the Acoustical Society of America (JASA)*, Vol. 67:2, pp. 648-662.
- Boë, L. J., Schwartz, J. L., Vallée, N., Abry, C., and Berrah, A., 1996, "La Matérialité des Structures Sonores du Langage 2 : de la Prédiction à l'Ontogenèse", in *Actes des XXIèmes Journées d'Etudes sur la Parole*, Avignon, pp. 143-146.
- Boersma, P., and Weenink, D., 2006, Praat. Doing Phonetics by Computer. version 4.4.34.

- Bohas, G., 1981, "Quelques Aspects de l'Argumentation et de l'Explication chez les Grammairiens Arabes", in *Arabica*, Vol. 28:2-3, pp. 204-221.
- Bohas, G., 1982, "Contribution à l'étude de la méthode des grammairiens arabes en morphologie et en phonologie : d'après des grammairiens arabes "tardifs"", Atelier national reprod., Université Lille 3, Lille, France.
- Bohas, G., 1985, "Explanation in Arab phonology", in *Journal of Arabic Linguistics*, Vol. 15, pp. 45-52.
- Bohas, G., 1993, "Développements récents en linguistique arabe et sémitique", Damas, Institut français de Damas.
- Bohas, G., and Guillaume, J.-P., 1984, "Étude des théories des grammairiens arabes", comprend : T.1, Morphologie et phonologie, Publications de l'Institut français de Damas.
- Bohas, G., Guillaume, J.-P., and Kouloughli, D. E., 1990, "The Arabic linguistic tradition", Routledge, London.
- Bomhard, A. S., 1995, "Historical-Comparative Linguistics", in *ORBIS*, N° 37, Louvain, pp. 3-59.
- Bonnot, J. F., 1976, "Contribution à l'étude des consonnes emphatiques de l'arabe à partir de méthodes expérimentales", Thèse de 3^{ème} cycle, Université de Strasbourg, Strasbourg, France.
- Bouzir, R., Nejmi, H., and Taki, M., 1991, "L'accent de l'arabe parlé à Casablanca et à Tunis : étude phonétique et phonologique", in *Proc. of the 12th International Congress of Phonetic Sciences (ICPhS)*, Aix-en-Provence, Vol. 5, pp. 134-137.
- Bradlow, A. R., 1995, "A comparative acoustic study of English and Spanish vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 97:3, pp. 1916-1924.
- Braham, A., 1997, "Al-tanzīm al-zamāniy fi al-^ṣarabiyya samā^ʿiyyan: dirāsa fil bunya al-kammiyya lil aṣwāt al-^ṣarabiyya", Thèse de Doctorat, Université de Tunis 1, Tunis, Tunisie.
- Brett, M., 1995, "The way of the nomad", in *Bulletin of the School of Oriental and African Studies (BSOAS)*, Vol. 58:2, pp. 251-269.
- Browman, C., and Goldstein, L., 1990, "Gestural specification using dynamically-defined articulatory structures", in *Journal of Phonetics*, Vol. 18, pp. 299-320.
- Browman, C., and Goldstein, L., 1992, "Articulatory Phonology: An Overview", in *Phonetica*, Vol. 49, pp. 155-180.
- Burgess, A., 1992, "The ear becomes the eye (Ch. 8)", in *A Mouthful of Air; Language, Languages...Especially English*, A. Burgess (Ed), William Morrow, New York, pp.72-93.
- Caelen-Haumont, G., 2005, "Les états émotionnels et la prosodie : paradigmes, modèles, paramètres (Ch. 15)", in *Phonologie et phonétique (Traité IC2, série Cognition et Traitement de l'Information)*, N. Nguyen, S. Wauquier-Gravelines, and J. Durand (Eds), Hermès, Paris, pp. 397-424.
- Calliope, 1989, "La parole et son traitement automatique", Masson (collection ENET et ENST).
- Cantineau, J., 1960, "Études de linguistique arabe", Librairie C. Klincksieck, Paris.
- Carré, R., 1996, "Prediction of vowel systems using a deductive approach", in *Proc. of International Congress on Speech and Language Processing*, Philadelphia, pp. 434-437.
- Carré, R., 1999, "Perception of coproduced speech gestures", in *Proc. of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Fransisco, pp. 643-646.
- Carré, R., 2004, "From an acoustic tube to speech production", in *Speech Communication*, Vol. 42:2, pp. 227-240.

- Carré, R., 2006, "Speech Gestures by Deduction, Gesture Production and Gesture Perception", in *Dynamics of Speech Production and Perception*, Proc. of the NATA Advanced Study Institute, IOS Press, Series I: Life and Behavioural Sciences, Vol. 374, P. Divenyi, S. Greenberg, and G. Meyer, pp. 59-70.
- Carré, R., 2007 (à paraître), "Dynamic properties of an acoustic tube: Prediction of vowel systems", in *Speech Communication*.
- Carré, R., and Chennoukh, S., 1995, "Vowel-Consonant-Vowel Modeling by Superposition of Consonant Closure on Vowel-to-Vowel Gestures", in *Journal of Phonetics*, Vol. 23:1-2, pp. 231-241.
- Carré, R., and Hombert, J.-M., 2002, "Variabilité phonétique en production et perception de la parole : stratégies individuelles", in *Invariants et Variabilité dans les Sciences Cognitives*, B. Mazoyer, J. Lautrey, and P. v. Geert (Eds), Presses de la Maison des Sciences de l'Homme, Paris, France, pp. 193-206.
- Carré, R., and Mrayati, M., 1992, "Distinctive regions in acoustic tubes. Speech production modelling", in *Journal d'Acoustique*, Vol. 5, pp. 141-159.
- Carré, R., and Mrayati, M., 1995, "Vowel transitions, vowel systems, and the Distinctive region Model", in *levels in Speech Communication*, C. Sorin, J. Mariani, H. Meloni, and J. Schoentgen (Eds), Elsevier, Amsterdam, pp. 73-89.
- Carré, R., Serniclaes, W., and Marsico, E., 2004, "Production and perception of vowel categories", in *Proc. of the meeting: From Sound to Sense*, MIT, pp. 1-6.
- Case, P., Tuller, B., Mingzhou, D., and Kelso, J. A. S., 1995, "Evaluation of a dynamical model of speech perception", in *Perception & Psychophysics*, Vol. 57:7, pp. 977-988.
- Caubet, D., 1983, "Quantification, Négation, Interrogation : les emplois de la particule si en arabe marocain", in *Arabica*, Vol. 30:3, pp. 227-245.
- Caubet, D., 1993, "*L'arabe marocain*", Peeters, Louvain/2 vol.
- Celdran, E. M., and Villalba, X., 1995, "Locus equation as a metrics for place of articulation in automatic speech recognition", in *Proc. of the 13th Internation Congress of Phonetic Sciences (ICPhS)*, Stockholm, Vol. 1, pp. 30-33.
- Chambers, J. K., Trudgill, P., and Schilling-Estes, N., 2002, "*The Handbook of Language Variation and Change*", Blackwell Publishers, Malden, MA.
- Chennoukh, S., Carré, R., and Lindblom, B., 1997, "Locus equations in the light of articulatory modeling", in *Journal of the Acoustical Society of America (JASA)*, Vol. 102:4, pp. 2380-2389.
- Chistovich, L. A., 1985, "Central auditory processing of peripheral vowel spectra", in *Journal of the Acoustical Society of America (JASA)*, Vol. 77:3, pp. 789-805.
- Cleveland, R. L., 1963, "A Classification for the Arabic Dialects of Jordan", in *Bulletin of the American Schools of Oriental Research*, Vol. 171, pp. 56-63.
- Cohen, D., 1962, "Koinè, langues communes et dialectes arabes", in *Arabica*, Vol. 9, pp. 119-144.
- Cohen, D., 1965, "Le système des voyelles brèves dans les dialectes maghrébins", in *Communications et rapports du 1^{er} Congrès International de Dialectologie Générale*, pp. 7-14.
- Cohen, D., 1968, "Les langues chamito-sémitiques", in *Le langage*, NRF-Gallimard, Paris, pp. 1288-1330.
- Cohen, D., 1972, "Problèmes de linguistique chamito-sémitique", in *Revue des Etudes Islamiques*, Vol. 40:1, pp. 43-68.
- Cohen, D., 1988, "Le chamito-sémitique", in *Les langues dans le monde ancien et moderne*, J. Perrot (Ed), CNRS, pp. 9-30.

- Cohen, M., 1924, "Langues chamito-sémitiques", in *Les Langues du Monde par un Groupe de Linguistes*, A. Meillet, and M. Cohen (Eds), hors-texte, planche 2, B., Champion, Paris, France, pp. 81-151,
- Colin, G. S., 1920, "Notes sur le parler arabe du nord de la région de Taza [avec 1 carte]", in *Le Bulletin de l'Institut français d'archéologie orientale (B.I.F.A.O.)*, Vol. 18, pp. 33-119.
- Colin, G.-S., 1975, "art. Al-Maghrib. VII. - Aperçu linguistique", in *Encyclopédie de l'Islam*, Leiden, p. 1193.
- Cooper, F. S., Delattre, P., Liberman, A. M., Borst, J. M., and Gerstman, L. J., 1952, "Some Experiments on the Perception of Synthetic Speech Sounds" in *Journal of the Acoustical Society of America (JASA)*, Vol. 24:6, pp. 597-606.
- Cowan, W., 1960, "Arabic evidence for proto-Semitic */awa/ and */o/", in *Language*, Vol. 36:1, pp. 60-62.
- Daniloff, R. G., and Hammarberg, R. E., 1973, "On defining coarticulation", in *Journal of Phonetics*, Vol. 1, pp. 239-248.
- de Jong, K. J., and Zawaydeh, B. A., 2002, "Comparing Stress, Lexical Focus, and Segmental Focus: Patterns of Variation in Arabic Vowel Duration", in *Journal of Phonetics*, Vol. 30:1, pp. 53-75.
- Delattre, P., 1966, "*Studies in French and Comparative Phonetics: selected papers in French and English*", The Hague Mouton & Co. (Ed), Paris, France.
- Delattre, P., Liberman, A. M., and Cooper, F. S., 1955, "Acoustic loci and transitional cues for consonants", in *Journal of the Acoustical Society of America (JASA)*, Vol. 27:4, pp. 769-773.
- Delattre, P., Liberman, A. M., Cooper, F. S., and Gerstman, L. J., 1952, "An experimental study of the acoustic determinants of vowel color; observations on one and two-formant vowels synthesized from spectrographic patterns", in *Word*, Vol. 8, pp. 195-210.
- Dell, F., and Elmedlaoui, M., 2002, "*Syllables in Tashlhiyt Berber and in Moroccan Arabic*", Kluwer Academic Publishers, Dordrecht, Boston, London.
- DiBenedetto, M.-G., 1989_a, "Vowel representation: Some observations on temporal and spectral properties of the first formant frequency", in *Journal of the Acoustical Society of America (JASA)*, Vol. 86:1, pp. 55-66.
- DiBenedetto, M.-G., 1989_b, "Frequency and time variations of the first formant: Properties relevant to the perception of vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 86:1, pp. 67-77.
- Disner, S. F., 1983, "Vowel quality: The relation between universal and language specific factors", in *UCLA Working Papers in Phonetics*, Vol. 58, pp. 1-158.
- Divenyi, P., Greenberg, S., and Meyer, G., 2006, "*Dynamics of Speech Production and Perception*", Proc. of the NATA Advanced Study Institute, IOS Press, Series I: Life and Behavioural Sciences, Vol. 374.
- Ehret, C., 1995, "*Reconstructing Proto-Afroasiatic (Proto-Afrasian). Vowels, Tone, Consonants, and Vocabulary*", Linguistics, 126, University of California Press.
- Elgendy, A. M., 2001, "*Aspects of pharyngeal coarticulation*", University of Amsterdam, LOT 44.
- Embarki, M., 1997, "La quantité vocalique en Arabe Marocain : entre l'apparement historique et la réalité acoustique", in *Actes des Journées d'Étude Linguistique : "La Voyelle dans tous ses états"*, Nantes, France, pp. 44-49.
- Embarki, M., 2007_a (à paraître), "Emergence et développement de la différenciation de genre entre 7 et 18 ans : perception de stimuli en arabe marocain en milieu scolaire au Maroc", in *Language and Gender in the Mediterranean, special issue of International Journal of the Sociology of Language*, M. Ennaji (Eds).

- Embarki, M., 2007_b, (à paraître), "Les dialectes arabes modernes : état et nouvelles perspectives pour la classification géo-sociologique," in *Arabica*.
- Embarki, M., and Guilleminot, C., 2003, "The Moving Boundaries of the First-Acquired Variety's Phonological Features: Evidence From Production/Perception of Moroccan Arabic's Vowels", in *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain, pp. 639-642.
- Engstrand, O., and Krull, D., 1991, "Effects of inventory size on the distribution of vowels in the formant space: Preliminary data from seven languages", in *PERILUS (Phonetic Experimental Research at the Institute of Linguistics University of Stockholm)*, Vol. 13, pp. 15-18.
- Ennaji, M., 2002_a, "Comment, Reviewer's comments: Diglossie", in *International Journal of the Sociology of Language*, Vol. 157, pp. 71-83.
- Ennaji, M., 2002_b, "Language contact, Arabization policy and education in Morocco", in *Language Contact and Language Conflict in Arabic. Variations on a Sociolinguistic Theme*, A. Rouchdy (Ed), Routledge Curzon, London, pp. 70-88.
- Fant, G., 1960, "*Acoustic theory of speech production*", Mouton, The Hague.
- Fant, G., 1973, "*Speech sounds and features*", MIT Press, Cambridge, MA.
- Ferguson, C., 1959, "Diglossia", in *Word*, Vol. 15, pp. 325-340.
- Fowler, C. A., 1980, "Coarticulation and theories of extrinsic timing", in *Journal of Phonetics*, Vol. 8, pp. 113-133.
- Fowler, C. A., 1986, "An event approach to the study of speech perception from a direct-realist perspective", in *Journal of Phonetics*, Vol. 14, pp. 3-28.
- Fowler, C. A., 1994, "Invariants, Specifiers, Cues: An Investigation of Locus Equations as Information for Place of Articulation", in *Perception & Psychophysics*, Vol. 55:6, pp. 597-610.
- Fowler, C. A., 1996, "Listeners do hear sounds, not tongues", in *Journal of the Acoustical Society of America (JASA)*, Vol. 99:3, pp. 1730-1741.
- Fry, D. B., 1966, "Mode de perception des sons du langage", in *Phonétique et Phonation*, A. Moles, and B. Vallancien (Eds), pp. 191-206.
- Fuchs, C., 2002, "Place et rôle de la variabilité dans les sciences cognitives", in *Invariants et variabilité dans les sciences cognitives*, J. Lautrey, B. Mazoyer, and P. van-Greet (Eds), Editions de la Maison des Sciences de l'Homme, Paris, pp. 157-173.
- Fujimura, O., and Ochiai, K., 1963, "Vowel identification and phonetic contexts", in *Journal of the Acoustical Society of America (JASA)*, Vol. 35:11, p. 1889A.
- Gay, T., 1977, "Articulatory Movements in VCV Sequences", in *Journal of the Acoustical Society of America (JASA)*, Vol. 62:1, pp. 183-193.
- Gay, T., 1978, "Effect of speaking rate on vowel formant movements" in *Journal of the Acoustical Society of America (JASA)*, Vol. 63:1, pp. 223-230.
- GEOATLAS, 2002, "GEOATLAS©", GRAPHI-OGRE, <http://www.geoatlas.fr/>
- Ghazali, S., 1977, "On the Controversy of the Arabic 'qāf'", in *Texas Linguistic Forum*, Vol. 6, pp. 40-50.
- Ghazali, S., 1979, "Du statut des voyelles en arabe", in *Analyse & Théorie*, Vol. 2/3, N° spécial : Etudes Arabes, pp. 199-219.

- Ghazali, S., Hamdi, R. and Barkat, M., 2002, "Speech Rhythm Variation in Arabic Dialects", in *Proc. of the 1st International conference on Speech Prosody*, Aix-en-Provence, France, pp. 331-334.
- Ghazeli, S., 1977, "*Back Consonants and Backing Coarticulation in Arabic*", Ph.D. Dissertation, University of Texas, Austin.
- Ghazeli, S., 1981, "La Coarticulation de l'emphase en arabe", in *Arabica*, Vol. 28: Juin-Sept., pp. 251-277.
- Girard, M., 2002, "*Étude et Modélisation de la production et de la perception de la parole*", Mémoire de DEA en Sciences Cognitives, Université Lyon 2, Lyon, France.
- Goldsmith, J. A., 1976_a, "*Autosegmental phonology*", Indiana University Linguistics Club, Bloomington.
- Goldsmith, J. A., 1976_b, "An overview of autosegmental phonology", in *Linguistic Analysis*, Vol. 2, pp. 23-68.
- Goldsmith, J. A., 1990 "*Autosegmental and Metrical Phonology*", Blackwell, Oxford.
- Gottfried, T. L., and Strange, W., 1980, "Identification of coarticulated vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 68:6, pp. 1626-1635.
- Greenberg, S., 1996, "Auditory Processing of Speech (Ch. 10)", in *Principles of Experimental Phonetics*, N. J. Lass (Ed), Mosby-Year Book, Inc., Missouri, pp. 362-407.
- Hagège, C., 1973, "*Profil d'un parler arabe du tchad*", Geuthner, Paris.
- Hamdi, R., 1991, "*Étude phonologique et expérimentale de l'emphase en arabe marocain de Casablanca*", Thèse de Doctorat en Sciences du Langage, Université Lyon 2, Lyon, France.
- Hardcastle, W. J., and Hewlett, N., 1999, "*Coarticulation: Theory, Data and Techniques*", Cambridge University Press.
- Harmegnies, B., and Poch-Olivé, D., 1992, "A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish", in *Speech Communication*, Vol. 11:4-5, pp. 429-437.
- Haroun, A., 1983, "*Al-Kitāb: Sibawayhi*", (5 volumes), 'ālam al-kutub, Beyrouth.
- Harris, C. M., 1953, "A Study of the Building Blocks in Speech", in *Journal of the Acoustical Society of America (JASA)*, Vol. 25:5, pp. 962-969.
- Harris, Z. S., 1942, "The Phonemes of Moroccan Arabic", in *Journal of the American Oriental Society*, Vol. 62:4, pp. 309-318.
- Hawkins, S., 1999_a, "Auditory Capacities and Phonological Development: Animal, Baby and Foreign Listeners (Ch. 13)", in *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory and Technology*, J.M. Pickett (Ed), Allyn & Bacon, Boston, pp. 183-197.
- Hawkins, S., 1999_b, "Looking for Invariant Correlates of Linguistic Units: Two Classical Theories of Speech Perception (Ch. 14)", in *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory and Technology*, J.M. Pickett (Ed), Allyn & Bacon, Boston, pp. 198-231.
- Hawkins, S., 1999_c, "Reevaluating Assumptions about Speech Perception: Interactive and Integrative Theories (Ch. 15)", in *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory and Technology*, J.M. Pickett (Ed), Allyn & Bacon, Boston, pp. 232-288.
- Hayward, R. J., 2000, "Afroasiatic, Chapter 4", in *African Languages, An Introduction*, B. Heine and D. Nurse (Eds), Cambridge University Press.

- Hetzron, R., 1974, "La division des langues sémitiques", in *Actes du Premier Congrès International de Linguistique Sémitique et Chamito-Sémitique*, Paris 16-19 Juillet 1969, A. Caquot & D. Cohen (Eds), The Hague Mouton, Paris, pp. 181-94.
- Hetzron, R., 1976, "Two principles of genetic classification", in *Lingua*, Vol. 38, pp. 89-108.
- Hetzron, R., 2001_a, "AfroAsiatic Languages (Ch. 8)", in *The Major Languages of South Asia, The Middle East and Africa*, B. Comrie (Ed), Routledge, pp. 151-159.
- Hetzron, R., 2001_b, "Semitic Languages (Ch. 9)", in *The Major Languages of South Asia, The Middle East and Africa*, B. Comrie (Ed), Routledge, pp. 160-169.
- Hillenbrand, J. M., Clark, M. J. and Nearey, T., 2001, "Effects of consonant environment on vowel formant patterns", in *Journal of the Acoustical Society of America (JASA)*, Vol. 109:2, pp. 748-762.
- Hillenbrand, J., and Gayvert, R. T., 1993, "Identification of steady-state vowels synthesized from the Peterson and Barney measurements", in *Journal of the Acoustical Society of America (JASA)*, Vol. 94:2 (part 1), pp. 668-674.
- Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K., 1995, "Acoustic characteristics of American English Vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 97:5 (Pt. 1), pp. 3099-3111.
- Hintze, J., 2004, "NCSS and PASS", Number Cruncher Statistical Systems. Kaysville, Utah, USA. <http://www.NCSS.com>
- Hombert, J.-M., 1984_a, "Phonétique expérimentale et diachronie : application à la tonogénèse", Thèse de Doctorat d'Etat, Université de Provence, Aix-en-Provence.
- Hombert, J.-M., 1984_b, "Réflexion sur le mécanisme des changements phonétiques", in *PHOLIA (PHONétique et LInguistique Africaine)*, Vol. 1, pp. 87-112.
- Hombert, J.-M., and Puech, G., 1984, "Espace vocalique et structuration perceptuelle : application au Swahili", in *PHOLIA (PHONétique et LInguistique Africaine)*, Vol. 1, pp. 199-208.
- Howell, D. C., 1998, "Méthodes statistiques en sciences humaines", De Boeck Université, Bruxelles.
- Hume, E., and Johnson, K., 2001, "The role of Speech Perception in Phonology", Academic Press, San Diego, California.
- Ibn-Jinni, A.-F., 10 siècle, "Sir Šinā'at Al-'i'rāb (les secrets du parler correct)", Vol. 1
- Ingham, B., 1971, "Some Characteristics of Meccan Speech", in *Bulletin of the Schools of Oriental and African Studies (BSOAS)*, Vol. 34:2, pp. 273-297.
- Ingham, B., 1973, "Urban and Rural Arabic in Khûzistân", in *Bulletin of the School of Oriental and African Studies (BSOAS)*, Vol. 36:3, pp. 533-553.
- Ingham, B., 1976, "Regional and social factors in the dialect geography of southern Iraq and Khûzistân", in *Bulletin of the School of Oriental and African Studies (BSOAS)*, Vol. 39:1, pp. 62-82.
- Iverson, P., and Kuhl, P., 1995, "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling", in *Journal of the Acoustical Society of America (JASA)*, Vol. 97:1, pp. 553-562.
- Iverson, P., and Kuhl, P., 1996, "Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/", in *Journal of the Acoustical Society of America (JASA)*, Vol. 99:2, pp. 1130-1140.
- Iverson, P., and Kuhl, P., 2000, "Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism?" in *Perception & Psychophysics*, Vol. 62:4, pp. 874-886.

- Jenkins, J. J., and Strange, W., 1999, "Perception of Dynamic Information for Vowels in Syllable Onsets and Offsets", in *Perception & Psychophysics*, Vol. 61:6, pp. 1200-1210.
- Jenkins, J. J., Strange, W., and Edman, T. R., 1983, "Identification of vowels in 'voiceless' syllables", in *Perception & Psychophysics*, Vol. 34, pp. 441-450.
- Jenkins, J. J., Strange, W., and Trent, S. A., 1999, "Context-Independent Dynamic Information for the Perception of Coarticulated Vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 106:1, pp. 438-448.
- Johnson, K., 1997, "*Acoustic and Auditory Phonetics*", Blackwell Publishers Inc., Oxford, UK.
- Johnson, K., 2000, "Adaptive dispersion in vowel perception", in *Phonetica*, Vol. 57:2-4, pp. 181-188.
- Johnson, K., and Mullenix, J. W., 1997, "*Talker Variability in Speech Processing*", Academic Press, San Diego, California.
- Johnson, K., Flemming, E., and Wright, R., 1993, "The hyperspace effect: Phonetic targets are hyperarticulated", in *Language*, Vol. 69, pp. 505-528.
- Johnson, K., Flemming, E., and Wright, R., 2004, "Letter to the Editor: Response to Whalen *et al.* 2004", in *Language*, Vol. 80:4, pp. 646-649.
- Johnson, K., Strand, E. A., and D'Imperio, M., 1999, "Auditory-visual integration of talker gender in vowel perception." in *Journal of Phonetics*, Vol. 27, pp. 359-384.
- Johnstone, T. M., 1967_b, "Aspects of syllabication in the spoken Arabic of 'Anaiza", in *Bulletin of the School of Oriental and African Studies (BSOAS)*, Vol. 30:1, pp. 1-16.
- Joos, M., 1948, "Acoustic Phonetics", in *Language: Journal of the Linguistic Society of America*, Vol. 24:2, Suppl..
- Kallas, E., 1999, "*Qui est arabophone ?*" Istituto di Sociologia Internazionale di Gorizia (I.S.I.G).
- Kaye, A. S., 2001, "Arabic, Chapter 10", in *The Major Languages of South Asia, The Middle East and Africa*, B. Comrie (Ed), Routledge, pp. 170-191.
- Kenstowicz, M., 1994, "*Phonology in Generative grammar*", Blackwell, Cambridge, MA.
- Kent, R., and Moll, K., 1969, "Vocal-Tract Characteristics of Stop Cognates", in *Journal of the Acoustical Society of America (JASA)*, Vol. 46:6 (part 2), pp. 1549-1555.
- Kewley-Port, D., 1982, "Measurements of formant transitions in naturally produced stop consonant-vowel syllables", in *Journal of the Acoustical Society of America (JASA)*, Vol. 72:2, pp. 379-389.
- Khomsi, A., 1975, "Étude Phonétique et Phonologique de l'Arabe Marocain de Casablanca". Thèse de 3^{ème} cycle, Université de Tours, Tour, France.
- Kingston, J., and Diehl, R. L., 1994, "Phonetic knowledge", in *Language*, Vol. 70, pp. 419-454.
- Klatt, D., and Klatt, L. C., 1990, "Analysis, synthesis, and perception of voice quality variations among male and female talkers", in *Journal of the Acoustical Society of America (JASA)*, Vol. 87:2, pp. 820-857.
- Koopmans-van-Beinum, F. J., 1980, "*Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions*", Dissertation. Academic Press, Amsterdam.
- Kozhevnikov, V. A., and Chistovich, L. A., 1965, "*Speech: Articulation and perception*", Joint Publications Research, Washington DC.
- Krull, D., 1987, "Second Formant Locus patterns as a measure of Consonant-Vowel Coarticulation", in *PERILUS (Phonetic Experimental Research at the Institute of Linguistics, University of Stockholm)*, Vol. 5, pp. 43-61.

- Krull, D., 1988, "Acoustic properties as predictors of perceptual responses: A study of Swedish voiced stops", in *PERILUS (Phonetic Experimental Research at the Institute of Linguistics, University of Stockholm)*, Vol. 7, pp. 66-70.
- Kuehn, D. P., and Moll, K., 1976, "A cineradiographic study of VC and CV articulatory velocities", in *Journal of Phonetics*, Vol. 4, pp. 303-320.
- Kuhl, P. K., 1991, "Human adults and Human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not", in *Perception & Psychophysics*, Vol. 50:2, pp. 93-107.
- Kuhl, P. K., 1994, "Learning and representation in speech and language". *Current Opinion in Neurobiology*, Vol. 4, pp. 812-822.
- Kuhl, P., Williams, K., Lacerda, F., Stevens, K., and Lindblom, B., 1992, "Linguistic experience alters phonetic perception in infants by 6 months age", in *Science*, Vol. 255, pp. 606-608.
- Ladefoged, P., 1967, *Three areas of experimental phonetics*, Oxford University Press, London.
- Ladefoged, P., 1996, *Elements of acoustic phonetics*, University of Chicago Press, Chicago.
- Ladefoged, P., and Broadbent, D. E., 1957, "Information conveyed by vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 29:1, pp. 98-104.
- Ladefoged, P., and Maddieson, I., 1996, *The sounds of the world's languages*, Blackwell Publisher, Oxford.
- Landercy, A., and Renard, R., 1977, *Éléments de phonétique*, Didier, Bruxelles.
- Laria, M., 1996, "Some characteristics features of Cyrenaican Arabic", in *Proc. of the 2nd International Conference of AIDA (Association Internationale de Dialectologie Arabe)*, pp. 123-132.
- Lass, N. J., 1996, *Principles of Experimental Phonetics*, N. J. Lass (Ed), Mosby-Year Book, Inc., Missouri.
- Lehiste, I., and Meltzer, D., 1973, "Vowel and speaker identification in natural and synthetic speech", in *Language and Speech*, Vol. 16, pp. 356-364.
- Lentin, J., 1994, "Classification et typologie des dialectes du BILAD AL-SHAM. Quelques suggestions pour un réexamen", in *Matériaux arabes et sudarabiques (M.A.S) du Groupe de Linguistique et de Littérature Arabes et Sudarabiques (G.E.L.L.A.S.)*, Paris, Vol. 6 (Numéro Spécial), pp. 11-44.
- Lieberman, A. M., 1970, "The grammars of speech and language", in *Cognitive Psychology*, Vol. 1, pp. 301-323.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studert-Kennedy, M., 1967_a, "Perception of the speech code", in *Psychological Review*, Vol. 74:6, pp. 431-461.
- Lieberman, A. M., Isenberg, D. S., and Rakerd, B., 1981, "Duplex perception of cues for stop consonants: Evidence for a phonetic mode", in *Perception & Psychophysics*, Vol. 30, pp. 133-143.
- Lieberman, A., 1957, "Some Results of Research on Speech Perception", in *Journal of the Acoustical Society of America (JASA)*, Vol. 29:1, pp. 117-123.
- Lieberman, A., and Mattingly, I., 1985, "The motor theory of speech perception revised", in *Cognition*, Vol. 21:1, pp. 1-36.
- Liljencrants, J., and Lindblom, B., 1972, "Numerical Simulation of Vowel Quality Systems: The role of Perceptual Contrast", in *Language*, Vol. 48, pp. 839-862.
- Lindblom, B., 1963_a, "Spectrographic study of vowel reduction", in *Journal of the Acoustical Society of America (JASA)*, Vol. 35:11, pp. 1773-1781.

- Lindblom, B., 1963_b, "On vowel reduction", Report #29, The Royal Institute of Technology, Speech Transmission Laboratory, Stockholm, Sweden.
- Lindblom, B., 1986, "Phonetic universals in vowel systems", in *Experimental Phonology*, J. J. Ohala, and J. J. Jaeger (Eds), Academic Press, New York, pp. 13-44.
- Lindblom, B., 1990, "Explaining phonetic variation: A sketch of the H and H theory", in *Speech production and speech modelling*, W. J. Hardcastle, and A. Marchal (Eds), Kluwer Academic Publishers, Dordrecht, pp. 403-439.
- Lindblom, B., 1996, "Role of Articulation in speech perception: Clues from production", in *Journal of the Acoustical Society of America (JASA)*, Vol. 99:3, pp. 1683-1692.
- Lindblom, B., and Studdert-Kennedy, M., 1967, "On the role of formant transitions in vowel recognition", in *Journal of the Acoustical Society of America (JASA)*, Vol. 42:4, pp. 830-843.
- Lindblom, B., Mauk, C., and Moon, S.-J., 2006, "Dynamic Specification in the Production of Speech and Sign", in *Dynamics of Speech Production and Perception*, Proc. of the NATA Advanced Study Institute, IOS Press, Series I: Life and Behavioural Sciences, Vol. 374, P. Divenyi, S. Greenberg, and G. Meyer, pp. 7-20.
- Livijn, P., 2000, "Acoustic distribution of vowels in differently sized inventories – hot spots or adaptive dispersion?" in *the Proc. of the XIIIth Swedish Phonetics Conference (FONETIK 2000)*, Skövde, Sweden, May 24-26, 2000, pp. 93-96.
- Lowenstamm, J., 1993, "Voyelles centrales, noyaux vides et statut de la longueur vocalique dans deux branches du sémitique", in *Linguistique arabe et sémitique, développements récents*, G. Bohas (Ed), Publication de l'Institut Français de Damas, pp. 77-91.
- Macchi, M., 1980, "Identification of vowels spoken in isolation versus vowels spoken in consonantal context", in *Journal of the Acoustical Society of America (JASA)*, Vol. 68:6, pp. 1636-1642.
- Mackenzie-Beck, J., 1997, "Organic variation of the vocal apparatus", in *The Handbook of Phonetic Sciences*, W. J. Hardcastle, and J. Laver (Eds), Blackwell, Oxford, GB, pp. 256-297.
- Maddieson, I., 1984, *Patterns of sounds*, Cambridge University Press, Cambridge.
- Maddieson, I., and Precoda, K., 1991, "Updating UPSID", in *UCLA Working Papers in Phonetics*, Vol. 74, pp. 104-111.
- Marçais, P., 1975, "Algérie. V — Langues", in *Encyclopédie de l'islam*, Brill et Maisonneuve, Leiden et Paris, pp. 384-390.
- Marçais, P., 1977, *Esquisse grammaticale de l'arabe maghrébin*, Adrien-Maisonneuve, Paris.
- Marçais, W., 1961, "Comment l'Afrique du Nord a été arabisée", in *William Marçais, Articles et Conférences*, Adrien-Maisonneuve, Paris, pp. 171-192.
- Martinet, A., 1953, "Remarques sur le consonantisme sémitique", in *Bulletin de la Société de Linguistique de Paris*, Vol. 49, 1, pp. 67-78.
- McCarthy, J., 1979, *Formal Problems in Semitic Phonology and Morphology*, Ph.D. Dissertation, MIT.
- McCarthy, J., 1981, "A prosodic theory of nonconcatenative morphology", in *Linguistic Inquiry*, Vol. 12:3, pp. 373-418.
- McCarthy, J., 1982, "Prosodic templates, morphemic templates and morphemic tiers", in *The Structure of Phonological Representations*, H. van der Hulst, and N. Smith (Eds), pp. 191-223.
- McCarthy, J., 1994, "The Phonetics and phonology of Semitic Pharyngeals", in *Phonological Structure and Phonetic Form: Papers in laboratory phonology 3*, P. A. Keating (Ed), Cambridge University Press, Cambridge, pp. 191-233.

- McCarthy, J., and Prince, A., 1988, "Prosodic morphology and templatic morphology", in *Perspectives on Arabic Linguistics, Current Issues in Linguistic Theory*, Koerner, E.F. Konrad, M. Eid, and J. McCarthy (Eds), John Benjamins, Amsterdam, pp. 1-54.
- McCarthy, J., and Prince, A., 1990, "Prosodic Morphology and Templatic Morphology. Papers from Second Annual Symposium on Arabic Linguistics" in *Perspectives on Arabic Linguistics: papers from the second symposium*, M. Eid, and J. McCarthy (Eds), Benjamins, Amsterdam, pp. 1-54.
- McDougall, K., 2006, "Dynamic Features of Speech and the Characterisation of Speakers: Towards a New Approach Using Formant Frequencies", in *International Journal of Speech, Language and the Law*, Vol. 13:1, pp. 89-126.
- McNeill, D., and Lindig, K., 1973, "The Perceptual reality of Phonem, Syllables, Words, and Sentences", in *Journal of Verbal Learning and Verbal Behaviour*, Vol. 12, pp. 419-430.
- Mehiri, A., 1973, "*Les théories grammaticales d'Ibn Jimi*", Publications de l'Université de Tunis, Tunis.
- Messaoudi, L., 2002, "Le parler ancien de Rabat face à l'urbanisation linguistique", in *Proc. of the 4th Conference of AIDA (Association Internationale de Dialectologie Arabe)*, Marrakech, Maroc, pp. 223-233.
- Meunier, C., 2005, "Invariants et Variabilité en Phonétique (Ch. 13)", in *Phonologie et phonétique (Traité IC2, série Cognition et Traitement de l'Information)*, N. Nguyen, S. Wauquier-Gravelines and J. Durand (Eds), Hermès, pp. 350-374.
- Meunier, C., Frenck-Mestre, C., Lelekov-Boissard, T., and Besnerais, M. L., 2003, "Production and perception of vowels: Does the density of the system play a role?" in *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain, pp. 723-726.
- Miller, G. A., 1951, "*Language and communication*", McGraw-Hill, New York.
- Miller, J., 1989, "Auditory-perceptual interpretation of the vowel", in *Journal of the Acoustical Society of America (JASA)*, Vol. 85:5, pp. 2114-2134.
- Modarresi, G., Sussman, H. M., Lindblom, B., and Burlingame, E., 2005, "Letter to the Editor: Locus Equation Encoding of Stop Place: Revisiting the Voicing/VOT Issue", in *Journal of Phonetics*, Vol. 33, pp. 101-113.
- Moon, S.-J., and Lindblöm, B., 1994, "Interaction between duration, context and speaking style in English stressed vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 96:1, pp. 40-55.
- Mrayati, M., Carré, R., and Guérin, B., 1988, "Distinctive region and modes: A new theory of speech production", in *Speech Communication*, Vol. 7, pp. 257-286.
- Nearey, T. M., 1989, "Static, Dynamic, and relational properties in vowel perception", in *Journal of the Acoustical Society of America (JASA)*, Vol. 85:5, pp. 2088-2113.
- Nearey, T. M., and Assmann, P. F., 1986, "Modeling the role of inherent spectral change in vowel identification", in *Journal of the Acoustical Society of America (JASA)*, Vol. 80:5, pp. 1297-1308.
- Nicolaidis, K., 2003, "Acoustic Variability of Vowels in Greek Spontaneous Speech", in *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain, pp. 3221-3224.
- Niedzielski, N., 1999, "The effect of social information on the perception of sociolinguistic variables", in *Journal of Language and Social Psychology*, Vol. 18:1, pp. 62-85.
- Ohala, J. J., 1993, "The phonetics of sound change", in *Historical Linguistics: Problems and Perspectives*, C. Jones (Ed), Longman, London, pp. 237-278.

- Palva, H., 1984, "A general classification for the Arabic dialects spoken in Palestine and Transjordan", in *Studia Orientalia*, Vol. 55:18, pp. 359-76.
- Parker, E. M., and Diehl, R. L., 1984, "Identifying vowels in CVC syllables: Effects of inserting silence and noise", in *Perception & Psychophysics*, Vol. 36, pp. 369-380.
- Peterson, G. E., 1952, "The information-bearing elements of speech", in *Journal of the Acoustical Society of America (JASA)*, Vol. 24:6, pp. 629-637.
- Peterson, G. E., and Barney, H. L., 1952, "Control methods used in a study of the vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 24:2, pp. 175-184.
- Pickett, J., 1999, *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Thechnology*, Allyn and Bacon, Boston.
- Pierrehumbert, J., 1992, "Dissimilarity in the Arabic Verbal Roots", in *Proc. of the 23rd Meeting of the Northeastern Linguistic Society*, Graduate Student Association, U. Mass. Amherst. pp. 367-381.
- Pisoni, D. B., and Remez, R. E., 2005, *The handbook of speech perception*, Blackwell Pub., Malden, MA.
- Piternann, M., 2000, "Effect of speaking rate and contrastive stress on formant dynamics and vowel perception", in *Journal of the Acoustical Society of America (JASA)*, Vol. 107:6, pp. 3425-3437.
- Pols, L., and van Son, R. J. J. H., 1993, "Acoustics and perception of dynamic vowel segments", in *Speech Communication*, Vol. 13, pp. 135-147.
- Puech, G., 1983, "Un fragment de phonologie polylectale", in *Principes de grammaire polylectale*, A. Berrendonner, M. Le-Guern, and G. Puech (Eds), P.U.L., Lyon, pp. 161-231.
- Rhardisse, N., 1995, *La coarticulation mandibulaire comme principe d'organisation syllabique en arabe de Fès et en Français*, Thèse de Doctorat Nouveau Régime en Sciences du Langage : Communication Parlée, Université Stendhal, Institut de la Communication Parlée, Grenoble, France.
- Rjaïbi, N., 1993, *Approches Historique, Phonologique et Acoustique de la Variabilité Dialectale Arabe : caractérisation de l'Origine Géographique en Arabe Standard.* Thèse de Doctorat Université de Franche-Comté, Besançon, France.
- Roman, A., 1981/82_a, "De la langue arabe comme un modèle général de la formation des langues sémitiques et de leur évolution", in *Arabica*, Vol. 28, pp. 127-61.
- Roman, A., 1981/82_b, "Le rôle des seuils de perception temporelle dans la constitution des contoïdes glottalisées", in *Travaux de l'Institut de Phonétique d'Aix*, Vol. 8, pp. 205-226.
- Rosch, E. H., 1975, "Cognitive reference points", in *Cognitive Psychology*, Vol. 7, pp. 532-547.
- Rosner, B. S., and Pickering, J. B., 1994, *Vowel perception and production*, Oxford University Press, Oxford [England]; New York.
- Roth, A., 1994, "Dialectologie comparée au Tchad et au Soudan. De quelques matériaux et discriminants", in *Matériaux arabes et sudarabiques (M.A.S) du Groupe de Linguistique et de Littérature Arabes et Sudarabiques (G.E.L.L.A.S.)*, Paris, Vol. 6, pp. 69-102.
- Ryalls, J. H., and Lieberman, P., 1982, "Fundamental frequency and vowel perception", in *Journal of the Acoustical Society of America (JASA)*, Vol. 72:5, pp. 1631-1634.
- Ryding, K. C., 2005, *A Reference Grammar of Modern Standard Arabic*, Cambridge University Press, Cambridge.

- Sabhi, N., 1997, "La variabilité dialectale arabe peut-elle être un moyen de reconnaissance de l'origine géographique ? Les fricatives interdentes, outils d'identification", in *Revue PARole*, Vol. 2, pp. 161-181.
- Sakarna, A. K., 1999, "*Phonological Aspects of 'Abady Arabic, A bedouin Jordanian Dialect*", Ph.D. Dissertation, University of Wisconsin-Madison.
- Sakarna, A. K., 2005, "The Linguistic Status of the Modern Jordanian Dialects", in *Arabica*, Vol. 52:4, pp. 522-543.
- Sanders, D. A., 1977, "*Auditory Perception of Speech*", Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
- Sandly, S., 2002, "L'Aspect Méthodologique de l'Atlas Linguistique de Tunisie", in *Proc. of the 5th Conference of AIDA (l'Association Internationale de Dialectologie Arabe)*, Cadiz, Espagne, pp. 445-457.
- Saussure, F. de, 1916, "*Cours de linguistique générale*", Payot, Paris (réédition, 1976).
- Schatz, C., 1954, "The role of context in the perception of stops", in *Language*, Vol. 30, pp. 47-56.
- Scholes, R. J., 1967, "Categorical responses to Synthetic Vocalic Stimuli by Speakers of Various Languages", in *Language and Speech*, Vol. 10:4, pp. 252-282.
- Scholes, R. J., 1967, "Phoneme categorization of synthetic vocalic stimuli by speakers of Japanese, Spanish, Persian and American English", in *Language and Speech*, Vol. 10, pp. 46-68.
- Scholes, R. J., 1968, "Perceptual categorisation of Synthetic Vowels as a tool in Dialectology and Typology", in *Language and Speech*, Vol. 11:3, pp. 194-207.
- Schroeder, M.R., Atal, B.S., and Hall, J.L., 1979, "Optimizing digital speech coders by exploiting masking properties of the human ear", in *Journal of the Acoustical Society of America (JASA)*, Vol. 66:6, pp. 1647-1652.
- Schwartz, J. L., Boë, L. J., Perrier, P., Guérin, B., and Escudier, P., 1989, "Perceptual Contrast and Stability in Vowel Systems: A 3-D Simulation Study", in *Proc. of EuroSpeech, 1*, pp. 63-66.
- Schwartz, J. L., Boë, L. J., Vallée, N. and Abry, C., 1997_b, "The Dispersion-Focalisation Theory of Vowel Systems", in *Journal of Phonetics*, Vol. 25, pp. 255-286.
- Schwartz, J. L., Boë, L. J., Vallée, N., and Abry, C., 1997_a, "Major trends in vowel system inventories", in *Journal of Phonetics*, Vol. 25, pp. 233-253.
- Sendlmeier, W. F., 1995, "Feature, phoneme, syllable or word: How is speech mentally represented?" in *Phonetica*, Vol. 52, pp. 131-143.
- Shahin, K. N., 1997, "Acoustics of pharyngealization vs. uvularization harmony", in *Current issues in linguistic theories: Perspectives on Arabic linguistics*, Vol. 153:10, pp. 215-237.
- Shahin, K. N., 2003, "Vowel innovation in Arabic: Inductive grounding and pattern symmetry", in *Research in Afroasiatic Grammar III (Selected papers of CAL 2000)*, In J. Lecarme, J. Lowenstamm, and Schlonsky (Eds), pp. 429-445.
- Sibawayhi, A. B. A., 8ème Siècle, "*Al-Kitāb*".
- Sjölander, K. 2002, "*The Snack sound toolkit (version 2.2)*", <http://www.speech.kth.se/snack/>.
- Slifka, J., 2003, "Tense/lax vowel classification using dynamic spectral cues", in *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain, pp. 921-924.
- SPSS© for Windows©, Rel. 11.0.0., 2001. Chicago: SPSS Inc. <http://www.spss.com/fr/>
- Stevens, K. N., 1960, "Toward a model for speech recognition", in *Journal of the Acoustical Society of America (JASA)*, Vol. 32:1, pp. 45-55.

- Stevens, K. N., 1972, "The Quantal Nature of Speech: Evidence from articulatory-acoustic data", in *Human communication: a unified view*, E. E. David, and P. B. Denes, (Eds), McGraw-Hill, New York, pp. 51–66.
- Stevens, K. N., 1989, "On the quantal nature of speech", in *Journal of Phonetics*, Vol. 17:1, pp. 3-45.
- Stevens, K. N., 1998, "*Acoustic Phonetics*", The MIT Press.
- Stevens, K. N., 2003, "Acoustic and Perceptual Evidence for Universal Phonological Features", in *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain, pp. 33-38.
- Stevens, K. N., and Keyser, S. J., 1989, "Primary features and their enhancements in consonants", in *Language*, Vol. 65, pp. 81-106.
- Stevens, K., and Blumstein, S. 1979, "Invariant cues for place of articulation in stop consonants", in *Journal of the Acoustical Society of America (JASA)*, Vol. 64:5, pp. 1358-1368.
- Stevens, K., Keyser, S. J., and Kawasaki, H., 1986, "Toward a phonetic and phonological theory of redundant features", in *Invariance and variability in speech processes*, J. Perkell, and D. Klatt (Eds), pp. 426-463.
- Strand, E. A., 1999, "Uncovering the role of gender stereotypes in speech perception", in *Journal of Language and Social Psychology*, Vol. 18:1, pp. 86-100.
- Strange, W., 1989, "Evolving Theories of Vowel Perception", in *Journal of the Acoustical Society of America (JASA)*, Vol. 85:5, pp. 2081-2087.
- Strange, W., 1999_a, "Perception of Vowels: Dynamic Constancy (Ch. 11)", in *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory and Technology*, J.M. Pickett (Ed), Allyn, and Bacon, Boston, pp. 153-165.
- Strange, W., 1999_b, "Perception of Consonants: From variance to invariance (Ch. 12)", in *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory and Technology*, J.M. Pickett (Ed), Allyn, and Bacon, Boston, pp. 166-182.
- Strange, W., and Bohn, O. S., 1998, "Dynamic Specification of Coarticulated German Vowels: Perceptual and Acoustical Studies", in *Journal of the Acoustical Society of America (JASA)*, Vol. 104:1, pp. 488-504.
- Strange, W., Jenkins, J. J., and Johnson, T. L., 1983, "Dynamic Specification of Coarticulated Vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 74:3, pp. 695-705.
- Strange, W., Verbrugge, R. R., Shankweiler, D. P., and Edman, T. R., 1976, "Consonant environment specifies vowel identity", in *Journal of the Acoustical Society of America (JASA)*, Vol. 60:1, pp. 213-224.
- Suleiman, Y., 1996, "*Language and Identity in the Middle East and North Africa*", Richmond Surrey: Curzon, pp. 81-99.
- Suleiman, Y., 2003, "*The Arabic Language and National Identity*", Georgetown University Press, Washington, D.C..
- Sussman, H. M., and Modarresi, G., 2003, "The Stability of Locus Equation Encoding of Stop Place", in *Proc. of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain, (3-9 September 2003), pp. 1931-1934.
- Sussman, H. M., and Shore, J., 1996, "Locus equations as phonetic descriptors of consonantal place of articulation", in *Perception & Psychophysics*, Vol. 58:6, pp. 936-946.
- Sussman, H. M., Bessell, N., Dalston, E., and Majors, T., 1997, "An investigation of stop place of articulation as a function of syllable position: A locus equation perspective", in *Journal of the Acoustical Society of America (JASA)*, Vol. 101:5, pp. 2826-2838.

- Sussman, H. M., Hoemeke, K. A., and Ahmed, F. S., 1993, "A Cross-Linguistic Investigation of Locus Equations as a Phonetic descriptor for Place of Articulation", in *Journal of the Acoustical Society of America (JASA)*, Vol. 94:3, pp. 1256-1268.
- Sussman, H. M., McCarthy, H. A., and Matthews, S. A., 1991, "An Investigation of Locus Equations as a source of relational invariance for stop place categorisation", in *Journal of the Acoustical Society of America (JASA)*, Vol. 90:3, pp. 1309-1325.
- Syrdal, A. K., and Gopal, H. S., 1986, "A perceptual model of vowel recognition based on the auditory representation of American English vowels", in *Journal of the Acoustical Society of America (JASA)*, Vol. 79:4, pp. 1086-1100.
- Tabain, M., 2000, "Coarticulation in CV Syllables: A Comparaison of Locus Equation and EPG Data", in *Journal of Phonetics*, Vol. 28, pp. 137-159.
- Tabain, M., and Butcher, A., 1999, "Stop consonants in Yanyuwa and Yindjibarndi: Locus Equation Data", in *Journal of Phonetics*, Vol. 27, pp. 333-357.
- Tatham, M., and Morton, K., 2006, "*Speech Production and Perception*", Palgrave Macmillan, July 2006.
- The MathWorks, (2002), "*MATLAB®: Matrix Laboratory, version 6.5*", <http://www.mathworks.fr/>.
- Thomas, E. R., 2002, "Instrumental Phonetics", in *The Handbook of Language Variation and Change*, J. K. Chambers, P. Trudgill, and N. Schilling-Estes (Eds), Blackwell Publishers, Malden, MA, pp. 168-200.
- Vaissière, J., 2006, "*La Phonétique*". Presses Universitaires de France (P.U.F.), collection: Que Sais-Je, Paris, 125 p.
- Vallée, N., 1994, "*Systèmes vocaliques : de la typologie aux prédictions*", Thèse de Doctorat en Sciences du Langage, Université Stendhal Grenoble, Grenoble, France.
- van Bergem, D. R., 1993, "Acoustic vowel reduction as a function of sentence accent, word stress, and word class", in *Speech Communication*, Vol. 12:1, pp. 1-23.
- van den Heuvel, H., Cranen, B., and Rietveld, T., 1996, "Speaker Variability in the Coarticulation of /a, i, u/", in *Speech Communication*, Vol. 18:2, pp. 113-130.
- Vanhove, M., 1996_a, "Les particules qad et ra' dans un dialecte arabe de la région de Yaafi' (Yémen)", in *Proc. of the 2nd International Conference of AIDA (Association Internationale de Dialectologie Arabe)*, Trinity Hall in the University of Cambridge, pp. 243-252.
- Vanhove, M., 2002, "Processus de koïnisation dans le dialecte arabe de Yafi'" in *Proc. of the 4th International Conference of AIDA (Association Internationale de Dialectologie Arabe)*, Marrakesh, Maroc, pp.151-159.
- Verbrugge, R. R., and Strange, W., 1976, "What information enables a listener to map a talker's vowel space?" in *Journal of the Acoustical Society of America (JASA)*, Vol. 60:1, pp. 198-212.
- Versteegh, K., 2001, "*The Arabic Language*", Edinburgh University Press, 2nd edition.
- Versteegh, K., 2006, "*Encyclopedia of arabic language and linguistics - Volume I, A-Ed*", K.Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski (Eds), Leiden, Boston, Brill Academic Publishers.
- Walter, H., 1982, "*Enquête phonologique et variétés régionales du français*", Presses Universitaires de France, Paris.
- Whalen, D. H., Magen, H. S., Pouplier, M., Kang, A. M., and Iskarous, K., 2004_a, "Vowel Production and Perception: Hyperarticulation without a Hyperspace Effect", in *Language and Speech*, Vol. 47:2, pp. 155-174.

- Whalen, D. H., Magen, H. S., Pouplier, M., Kang, A. M., and Iskarous, K., 2004_b, "letters to *Language*: Vowel targets without a hyperspace effect", in *Language*, Vol. 80:3, pp. 377-378.
- Willis, C., 1971, "Synthetic Vowel Categorisation and Dialectology", in *Language and Speech*, Vol. 14:3, pp. 213-228.
- Yéou, M., 1995, "*Une étude expérimentale des consonnes postérieures et pharyngalisées de l'arabe standard*", Thèse de Doctorat, Université Paris III, Paris, France.
- Yéou, M., 1997, "Locus Equations and the Degree of Coarticulation of Arabic Consonants", in *Phonetica*, Vol. 54:3-4, pp. 187-202.
- Younes, M., 1984, "On vowel shortening in Palestinian Arabic", in *Linguistic Theory 130: Perspectives on Arabic Linguistics VII*, Cornell University, pp. 157-171.
- Znagui, I., 1992, "Étude Formantique des voyelles de l'Arabe Standard", in *Actes des XIX^{èmes} Journées d'Étude sur la Parole*, Bruxelles (19-22/05/1992), pp. 493-497.
- Znagui, I., 1993, "Classification phonétique des voyelles de l'arabe standard", in *Colloque Phonologie/Phonétique de l'Université de York*, Toronto, Canada, Editions Mélodies, pp. 87-97.
- Znagui, I., 1995, "*Études Phonétique et Perceptive des voyelles de l'Arabe Standard Moderne*", Thèse de Doctorat en Sciences du Langage, Université Paris III, Paris, France.
- Znagui, I., and Bouchnafi, A., 1994, "L'écart entre F1 et F2 comme Mesure de l'Effet des Consonnes Linguales sur les Cibles Acoustiques des Voyelles de l'Arabe Standard Moderne", in *Actes des XX^{èmes} Journées d'Étude sur la Parole*, Trégastel, pp. 215-220.
- Znagui, I., and Boudelaa, S., 1993, "Towards an acoustic-phonetic classification of the modern standard arabic vowels", in *EuroSpeech*, Berlin, pp. 1427-1429.
- Znagui, I., and Yéou, M., 1996, "Le Rôle de la Coarticulation dans la Perception des Voyelles de l'Arabe Standard Moderne", in *Actes des XXI^{èmes} Journées d'Étude sur la Parole*, Avignon, pp. 35-38.

BIBLIOGRAPHIES CITÉES

- *Abbs, J. H., and Sussman, H. M., 1971, "Neurophysiological feature detectors and speech perception. A discussion of theoretical implications", in *Journal of Speech and Hearing Research*, Vol. 14, pp. 23-36. Cité dans Sanders (1977).
- *Abdel-Jawad, H. R., 1981, "*Lexical and Phonological Variation in Spoken Arabic in Amman*", Ph.D. Dissertation, University of Pennsylvania. Cité dans Sakarna (2005).
- *Abdel-Jawad, H. R., 1986, "The emergence of an urban dialect in the Jordaian urban centers", in *International Journal of the Sociology of Language*, Vol. 5, pp. 53-63. Cité dans Sakarna (2005).
- *Abul-Fadl, F., 1961, "*Volkstümliche Texte in arabischen Bauerndialekten der ägyptischen Provinz Šarqīyya mit dialektgeographischen Untersuchungen zur Lautlehr*", University of Münster. Cité dans Versteegh (2001).
- *Aguadé, J., and Benyahia, L., 1990, "El sáwes Bu-Azza visita al siquiatra. Un diálogo en árabe dialectal de Casablanca." in *Anaquel de Estudios Arabes*, Vol. 1, pp. 103-128. Cité dans Barkat (2000).
- *Ali, L., and Danilof, R. G., 1972, "A contrastive cinefluorographic investigation of the articulation of emphatic/non emphatic cognate consonants", in *Studia Linguistica*, Vol. 25:2, pp. 81-105. Cité dans Sakarna (1999).

- *Arnaud, V., and Konopckzynski, G., 1997, "Un glissement articulatoire au sein du système vocalique français : le cas de la variante haut-jurassienne", in *Actes des Journées d'Études Linguistiques « la voyelle dans tous ses états »*, Nantes, pp. 56-61. Cité par Meunier (2005).
- *Awwad, M. A., 1986, "Free and Bound Pronouns as Verbs in Rural Palestinian Colloquial Arabic", in *Zeitschrift der Deutschen Morgenlandischen Gesellschaft*, Vol. 137:2, pp. 297-331. Cité dans Barkat (2000).
- *Baccouche, T., and Mejri, S., 2000_a, "Atlas linguistique de Tunisie : identité et développement [en arabe]", in *Revue tunisienne des sciences sociales, numéro spécial : Langage et altérité : l'expérience de l'atlas linguistique de Tunisie*, CERES, Tunis, pp.17-24. Cité dans Sandly (2002).
- *Baccouche, T., and Mejri, S., 2000_b, "Pour un atlas linguistique arabe : la contribution tunisienne [en arabe]", in *Revue tunisienne des sciences sociales, numéro spécial : Langage et altérité : l'expérience de l'atlas linguistique de Tunisie*, CERES, Tunis, pp.25-34. Cité dans Sandly (2002).
- *Baccouche, T., and Mejri, S., 2000_c, "L'atlas linguistique de Tunisie : problématique phonologique", in *Revue tunisienne des sciences sociales, numéro spécial : Langage et altérité : l'expérience de l'atlas linguistique de Tunisie*, CERES, Tunis, pp.151-156. Cité dans Sandly (2002).
- *Baccouche, T., and Mejri, S., 2000_d, "L'atlas linguistique de Tunisie : spécificités phonologiques", in *Revue tunisienne des sciences sociales, numéro spécial : Langage et altérité : l'expérience de l'atlas linguistique de Tunisie*, CERES, Tunis, pp.157-162. Cité dans Sandly (2002).
- *Bani-Yassin, R., and Owens, J., 1984, "The Bduul Dialect of Jordan", in *Anthropological Linguistics*, Vol. 26:2, pp. 202-232. Cité dans Sakarna (1999).
- *Barrett, S. E., 1998, "*Prototypes in Speech Perception*", Unpublished Ph.D. Dissertation, University of Cambridge, Cambridge. Cité dans Hawkins (1999_c).
- *Barthélemy, A., 1935-69, "*Dictionnaire arabe-français, dialectes de Syrie : Alep, Damas, Liban, Jérusalem*", Institut de France, Paris. Cité dans Versteegh (2001).
- *Behnstedt, P., 1985, "*Die Nordjemenitischen Dialekte - Vol. I: Atlas*", L. Reichert (Ed), Wiesbaden. Cité dans Versteegh (2001).
- *Behnstedt, P., 1992, "*Die Nordjemenitischen Dialekte - Vol. II: Glossar (Buchstaben Alif-Dal)*", L. Reichert (Ed), Wiesbaden. Cité dans Versteegh (2001).
- *Behnstedt, P., and Woidich, M., 1985, "*Die Ägyptisch-Arabischen Dialekte. Vol. 1: Einleitung und Anmerkungen zu den Texten; Vol. II: Sprachatlas von Ägypten*", L. Reichert (Ed), Wiesbaden. Cité dans Versteegh (2001).
- *Behnstedt, P., and Woidich, M., 1987, "*Die ägyptisch-arabischen Dialekte: Vol. III, Texte: I. Delta Dialekte*", L. Reichert (Ed), Wiesbaden. Cité dans Versteegh (2001).
- *Behnstedt, P., and Woidich, M., 1988, "*Die ägyptisch-arabischen Dialekte: Vol. III, Texte: II. Niltaldialekte, III. Oasendialekte*", L. Reichert (Ed) Wiesbaden. Cité dans Versteegh (2001).
- *Benkirane, T., 1998, "Intonation in Western Arabic (Morocco)", in *Intonation Systems, a Survey of Twenty Languages*, D. Hirst, and A. Di-Cristo (Eds), Cambridge University Press, pp. 345-359. Cité dans Benkirane (2000).
- *Bergsträßer, G., 1915, "Sprachatlas von Syrien and Palästina (+ maps XX-LXII)", in *Zeitschrift des Deutschen Palästina-Vereins (ZDPV)*, Vol. 38, pp. 168-222. Cité dans Versteegh (2001).
- *Berry, M. F., 1969, "*Language Disorders of Children*", Appleton-Century-Crofts, New York. Cité dans Sanders (1977).

- *Blanc, H., 1964, "*Communal Dialects in Baghdad*", Cambridge. Cité dans Versteegh (2001).
- *Borg, A., 1984, "Kormatiki Arabic: Phonological Notes and Texts", in *Zeitschrift für Arabische Linguistik (Z.A.L.)*, Vol. 13, pp. 68-85. Cité dans Barkat (2000).
- *Brockelmann, C., (1910), "*Précis de linguistique sémitique*". Librairie C. Klincksieck, Paris. Cité dans Barkat (2000).
- *Cadora, F. J., 1992, "*Bedouin, Village, and Urban Arabic: An Ecolinguistic Study*", E.J. Brill (Ed), Leiden. Cité dans Versteegh (2001).
- *Cantineau, J., 1935, "*Grammaire du Palmyrénien épigraphique*", Publications de l'Institut d'Etudes Orientales de la faculté des lettres d'Alger, 4. Cairo. Cité dans Versteegh (2001).
- *Cantineau, J., 1936, "Études sur quelques parlers de nomades arabes d'Orient", in *Annales de l'Institut d'Etudes Orientales, Algiers*, Vol. 2, pp. 1-118. Cité dans Versteegh (2001).
- *Cantineau, J., 1937, "Études sur quelques parlers de nomades arabes d'Orient", in *Annales de l'Institut d'Etudes Orientales, Algiers*, Vol. 3, pp. 119-237. Cité dans Versteegh (2001).
- *Cantineau, J., 1938, "Remarques sur les parlers de sédentaires syro-libano sédentaires", in *Bulletin de la Société de Linguistique de Paris*, Vol. 40, pp. 80-89. Cité dans Versteegh (2001).
- *Cantineau, J., 1955, "La dialectologie arabe", in *Orbis*, Vol. 4, pp. 149-169. Cité dans Cantineau (1960).
- *Carter, M. G. 1981, "*Arab Linguistics: An Introductory Classical Text with Translation and Notes*", Benjamins, Amsterdam. Cité dans Versteegh (2001).
- *Caubet, D., 1998, "Étude sociolinguistique des traits préhilaliens dans un dialecte en voie d'urbanisation à Fès", in *Peuplement et Arabisation au Maghreb Occidental. Dialectologie et Histoire*, J. Aguadé, P. Cressier, and A. Vicente (Eds), Madrid-Zaragoza : coéditions de la Casa de Velazquez, pp. 165-175. Cité dans Messaoudi (2002).
- *Cohen, D., 1963, "*Le dialecte arabe hassaniya de Mauritanie*", Klincksieck, Paris. Cité dans Versteegh (2001).
- *Cohen, D., 1964, "*Le parler arabe des Juifs de Tunis. Vol. 1 : textes et documents linguistiques et ethnographiques*", Mouton, The Hague - Paris. Cité dans Versteegh (2001).
- *Cohen, D., 1973, "Pour un atlas linguistique et sociolinguistique de l'arabe", in *Actes du 1er Congrès d'étude des cultures méditerranéennes d'influence arabo-berbère*, M. Galley (Ed), SNED, Alger, pp. 63-69. Cité dans Barkat (2000).
- *Cohen, D., 1975, "*Le parler arabe des Juifs de Tunis. Tome II : Étude linguistique*", Mouton, The Hague - Paris. Cité dans Versteegh (2001).
- *Cohen, M., 1912, "*Le parler arabe des Juifs d'Alger*", Champion, Paris. Cité dans Barkat (2000).
- *Cooper, F. S., 1972, "How is language conveyed by speech?" in *Language by ear and by eye: The relationships between speech and reading*, J. F. Kavanagh, and I. G. Mattingly (Eds), MIT Press, Cambridge, MA, pp. 25-45. Cité dans Sanders (1977).
- *Czapkiewicz, A., 1960, "*Sprachproben aus Mādabā*", Polska Akademia Nauk, Krakow. Cité dans Barkat (2000).
- *Daniloff, R. G., 1973, "Normal articulation processes", in *Normal Aspects of Speech, Hearing, and Language*, F. D. Minifie, T. J. Hixon, and F. Williams (Eds), Prentice-Hall, Englewood cliffs, New Jersey. Cité dans Sanders (1977).
- *Diehl, R. L., Kluender, K. R., and Walsh, M. A., 1990, "Some auditory bases of speech production and perception", in *Advances in Speech, Hearing and Language Processing*, W. A. Ainsworth (Ed), JAI Press, London, pp. 243-268. Cité dans Hawkins (1999_b).

- *Diem, W., 1978, "Divergenz und Konvergenz im Arabischen", in *Arabica*, Vol. 25, pp. 128-147. Cité dans Versteegh (2001).
- *Elbiad, M., 1985, "*A Sociolinguistic Study of the Arabization Process and Its Conditioning Factors in Morocco*", State University of New York, Buffalo. Cité dans Ennaji (2002_b).
- *Ennaji, M., 1988, "Language Planning in Morocco and Changes in Arabic", in *International Journal of the Sociology of Language*, Vol. 74, pp. 9-39. Cité dans Ennaji (2002_a).
- *Ennaji, M., 1991, "Aspects of Multilingualism in the Maghreb", in *International Journal of the Sociology of Language*, Vol. 87, Sociolinguistics of the Maghreb, pp. 7-25. Cité dans Ennaji (2002_a).
- *Ennaji, M., and Sadiqi, F., 1994, "*Applications of Modern Linguistics : Casablanca : Afrique-Orient*". Cité dans Ennaji (2002_a).
- *Fairbanks, G., and Grubb, P. A., 1961, "A psychophysical investigation of vowel formants", in *Journal of Speech and Hearing Research*, Vol. 4, pp. 203-219. Cité dans Strange *et al.* (1976).
- *Fant, G., 1962, "Descriptive analysis of the acoustic aspects of speech", in *Logos*, Vol. 5:1, pp. 3-17. Cité dans Sanders (1977).
- *Fant, G., 1967, "Auditory Patterns of Speech", in *Models for the Perception of Speech and Visual Form*, Watten-Dunn (Ed), The MIT Press, Cambridge, pp. 111-125. Cité dans Sanders (1977).
- *Ferguson, C., 1956, "Two Problems in Arabic Phonology", in *Word*, Vol. 13:3, pp. 460-78. Cité dans Barkat (2000).
- *Fischer, W., 1967, "Silben struktur und Vokalismus im Arabischen", in *Zeitschrift der Deutschen Morgenländischen Gesellschaft (ZDMG)*, Vol. 117, pp. 30-77. Cité dans Barkat (2000).
- *Fischer, W., 1982, "Grundriss der arabischen Philologie, I: Sprachwissenschaft", Reichert, Wiesbaden. Cité dans Barkat (2000).
- *Fischer, W., and Jastrow, O., 1980, "*Handbuch der arabischen Dialekte*", Otto Harrassowitz, Wiesbaden. Cité dans Versteegh (2001).
- *Fleisch, H., 1974, "*Etudes d'arabe dialectal*", Dar El-Machreq, Beyrouth. Cité dans Versteegh (2001).
- *Fowler, C. A., and Rosenblum, L. D., 1991, "The perception of phonetic gestures", in *In Modularity and the Motor Theory of Speech Perception, Proc. of a Conference in the Honor of Alvin M. Liberman*, I. G. Mattingly and M. S.-K. (Eds.), Hillsdale, NJ: Erlbaum, pp. 33-59. Cité dans Hawkins (1999_b).
- *Fujisaki, H., and Kawashima, T., 1968, "The role of pitch and higher formants in the perception of vowels", in *IEEE Transmission, Audio, and Electroacoustics*, Vol. AU-16, pp. 73-77. Cité dans Strange (1999_a).
- *Gibson, J. J., 1954, "A theory of pictorial perception", in *Audio-Visual Communication Review*, Vol. 1, pp. 3-23. Cité dans Tatham and Morton (2006).
- *Gibson, J. J., 1966, "*The senses Considered as Perceptual Systems*", Houghton Mifflin, Boston. Cité dans Hawkins (1999_c).
- *Grand'Henry, J., 1972, "*Le parler arabe de Cherchell (Algérie)*", Louvain-La-Neuve. Cité dans Versteegh (2001).
- *Herzallah, R. S., 1990, "*Aspects of Palestinian Arabic Phonology: A Nonlinear Approach*", Ph.D. Dissertation at the Dept. of Modern Languages and Linguistics at Cornell University. Cité dans Sakarna (1999).
- *Holes, C., 1990, "*Gulf Arabic*", London and New York: Routledge. Cité dans Versteegh (2001).

- *Honikman, B., 1964, "Articulatory settings", in *Honour of Daniel Jones*, D. Abercrombie, D. Fry, P. MacCarthy, N. Scott, and J. Trim (Eds), Longmans, London, pp. 73-84. Cité dans Bradlow (1995).
- *Huang, C. B., 1985, "*Perceptual Correlates of the Tense/Lax Distinction in General American English*", MIT, Cambridge, MA. Cité dans Strange (1999_a).
- *Hussein, R. F., and El-Ali, N., 1989, "Subjective Reactions of Rural University Students toward different varieties of Arabic", in *Al-ʿArabiyya (Journal of the American Association of Teachers of Arabic)*, Vol. 22, pp. 37-54. Cité dans Sakarna (2005).
- *Ibn-Khaldoun, A. Z. A., 14^{ème} siècle, "Kitāb Al-ʿibar, wa Diwān Al-Mubtada wa'l Ḥabar, fi Ayyami'l ʿArab wa'l ʿağm wa'l Barbar" (Livre des Exemples Instructifs et Recueil d'Origines et de Récits, concernant l'Histoire des Arabes, des Peuples Etrangers et des Berbères), Vol. 3 : "Histoire des Berbères et des Dynasties Musulmanes de l'Afrique Septentrionale, traduit de l'arabe par W. Mac-Guckin de Slane", Berti Editions, Alger (2003). Cité dans Embarki (2007_b, à paraître).
- *Ingham, B., 1982, "*North east Arabian dialects*", Kegan Paul International, London and Boston. Cité dans Versteegh (2001).
- *Iriksoussi, B., 1981, "*La Structure Acoustique et Phonétique des Consonnes Arabes dans le Parler de Damas*", Thèse de 3^{ème} Cycle, Université de Provence, Aix-en-Provence, France. Cité dans Barkat (2000).
- *Irsheid, O. M., and Kenstowicz, M., 1984, "Some phonological rules of Bani-Hassan Arabic", in *Studies in the Linguistics Sciences*, Vol. 14:1, pp. 109-47. Cité dans Sakarna (1999).
- *Irshied, O.M., 1984, "*The Phonology of Arabic; Bani Hassan, A Bedouin Jordanian Dialect*", Ph.D. Dissertation at the University of Illinois, Urbana-Champaign. Cité dans Sakarna (2005).
- *Jastrow, O., 1978, "*Die mesopotamisch-arabischen qeltu-Dialekte. Vol. I. Phonologie und Morphologie*", F. Steiner (Ed), Wiesbaden. Cité dans Versteegh (2001).
- *Jastrow, O., 1979, "Der arabischen Mundart von Mossul", in *Zeitschrift für Arabische Linguistik (Z.A.L)*, Vol. 2, pp. 36-75. Cité dans Versteegh (2001).
- *Jastrow, O., 1981, "*Die mesopotamisch-arabischen qeltu-Dialekte. Vol. II. Texte*", F. Steiner (Ed), Wiesbaden. Cité dans Versteegh (2001).
- *Jastrow, O., 1990, "*Der arabische Dialekt der Juden von 'Aqra und Arbīl*", O. Harrassowitz (Ed), Wiesbaden. Cité dans Versteegh (2001).
- *Jastrow, O., 1994, "The qeltu Arabic dialects of Mesopotamian Arabic", in *Actas del Congreso Internacional sobre Interferencias Lingüísticas Arabo-Romances y Paralelos Extra-Ibéricos*, Navarro, and Navarro, Zaragoza, pp. 119-123. Cité dans Versteegh (2001).
- *Johnstone, T. M., 1967_a, "*Eastern Arabian Dialect Studies*", London. Cité dans Versteegh (2001).
- *Jongman, A., Fourakis, M., and Sereno, J., 1989, "The acoustic vowel space of Modern Greek and German", in *Language and Speech*, Vol. 32, pp. 221-248. Cité dans Bradlow (1995).
- *Kaye, A. S., 1976, "*Chadian and Sudanese Arabic in the light of comparative arabic Dialectology*", Mouton, The Hague. Cité dans Versteegh (2001).
- *Kaye, A. S., 1982, "*A Dictionary of Nigerian Arabic*", Undena Publications, Malibu. Cité dans Versteegh (2001).
- *Kaye, A. S., 1986, "*Nigerian Arabic-English Dictionary*", Undena Publications, Malibu. Cité dans Barkat (2000).

- *Kaye, A. S., and Daniel, P. T., 1997, "*Phonologies of Asia and Africa (including the Caucasus)*". In A.S. Kaye (Ed), Eisenbrauns, Winona Lake. Cité dans Barkat (2000).
- *Kuhl, P., 1993, "Innate predispositions and the effects of experience in speech perception: The native language magnet theory", in *Developmental neurocognition: Speech and face processing in the first year of life*, I. B. de-Boysson-Bardies, S. d. Schonen, P. Jusczyk, P. McNeilage, and J. Morton (Eds), Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 259-274. Cité dans Iverson & Kuhl (2000).
- *Kuhl, P., and Iverson, P., 1995, "Linguistic experience and the "perceptual magnet effect"", in *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research*, Timonium, MD: York Press Inc. pp. 121-154. Cité dans Hawkins (1999_b).
- *Laabi-Zemmama, T., 1975, "*Phonétique et phonologie du parler arabe de Fès (Maroc)*", Thèse 3^{ème} Cycle, Université Toulouse 2, Toulouse, France. Cité dans Rhardisse (1995).
- *Labov, W., Yaeger, M. and Steiner, R., 1972, "*A quantitative study of sound change in progress*", US Regional Survey, Philadelphia. Cité dans Thomas (2002).
- *Lentin, J., 1995-96, "Existait-il un "arabe levantin" à l'époque ottomane ? Traits dialectaux communs et koine para-standard transrégionale", in *Matériaux arabes et sudarabiques (M.A.S) du Groupe de Linguistique et de Littérature Arabes et Sudarabiques (G.E.L.L.A.S.)*, Paris, Vol. 6 (Numéro Spécial), pp. 99-113. Cité dans Barkat (2000).
- *Lethem, G. L., 1920, "*Colloquial Arabic, Shuwa dialect of Bornu, Nigeria and of the region of Lake Tchad*", The Crown Agents for the Colonies, London. Cité dans Barkat (2000).
- *Levi-Provençal, E., 1922, "*Textes arabes de l'Ouargha. Dialecte des Jbala (Maroc septentrional)*", Leroux, Paris. Cité dans Messaoudi (2002).
- *Lévy, S., 1995, "Problèmes de géographie dialectale : strates et buttes témoins (l'exemple du parler juif de Tafilalt)", in *Dialectologie et sciences humaines au Maroc, Publications de la faculté des lettres de Rabat*, pp. 51-59. Cité dans Messaoudi (2002).
- *Liberman, A. M., Cooper, F. S., Harris, K. S., MacNeilage, P. F., and Studdert-Kennedy, M., 1967_b, "Some observations on a model for speech perception", in *Models for the perception of speech and visual form*, W. Wathen-Dunn (Ed), MIT Press, Cambridge, MA, pp. 68-87. Cité dans Sanders (1977).
- *Liberman, A. M., Mattingly, I. G., and Turvey, M. T., 1972, "Language codes and memory codes (Ch. 13)", in *Coding processes in human memory*, A. W. Melton, and E. Martin (Eds), Washington, DC, pp. 307-334. Cité dans Sanders (1977).
- *MacKay, D. M., 1956, "The Epistemological Problem for Automata". In *Automata Studies*, C. E. Shannon and J. McCarthy, (Eds), Princeton University Press, pp. 235-251. Cité dans Sanders (1977).
- *MacNeilage, P. F., 1970, "Motor control of serial ordering of speech", in *Psychological Review*, Vol. 77, pp. 182-196. Cité dans Strange (1989).
- *Marçais, P., 1956, "*Le parler arabe de Djidjelli (Nord Constantinois, Algérie)*", Lib. Adrien-Maisonneuve, Paris, 648 p. Cité dans Versteegh (2001).
- *Marçais, W., 1902, "*Le dialecte arabe parlé à Tlemcen : grammaire, textes et glossaire*", Leroux, Paris. Cité dans Versteegh (2001).

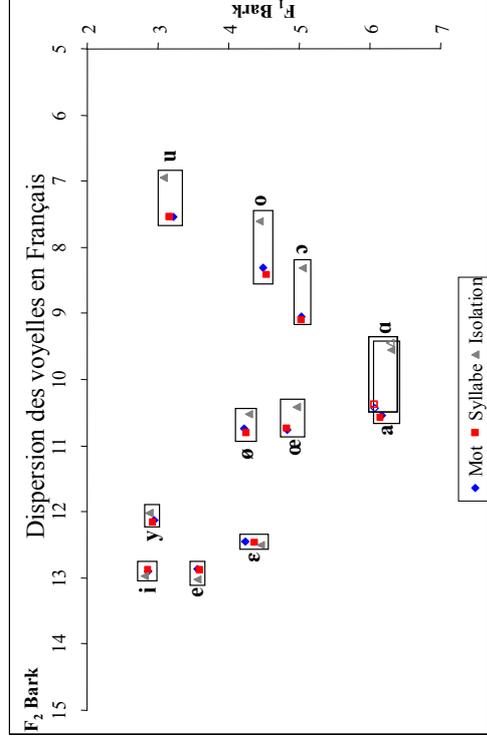
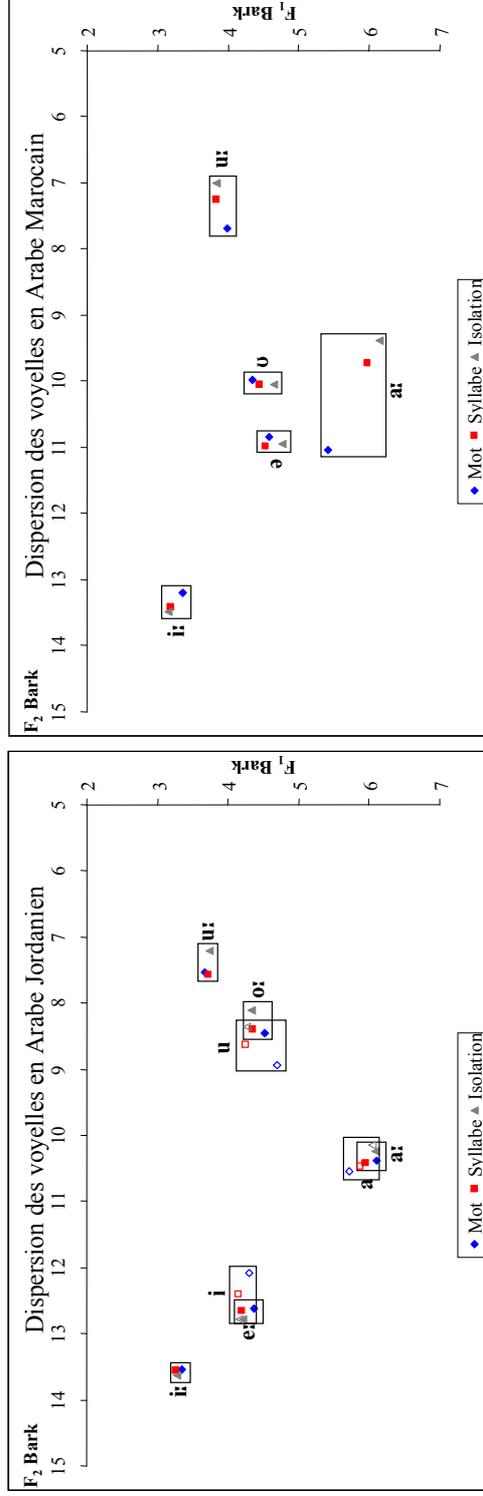
- *Marçais, W., 1930-31, "La diglossie : un pèlerinage aux sources", in *Bulletin de la Société de Linguistique de Paris*, Vol. 76:1, pp. 61-98. Cité dans Ennaji (2002_a).
- *Mattingly, I. G., 1972, "Speech cues and sign stimuli: An ethological view of speech perception and the origin of language", in *American Scientist*, Vol. 60, pp. 327-337. Cité dans Sanders (1977).
- *McGurk, H., and MacDonald, J., 1976, "Hearing lips and seeing voices", in *Nature*, Vol. 264, pp. 746-748. Cité dans Thomas (2002).
- *Owens, J., 1984, "*A Short Reference Grammar of Eastern Libyan Arabic*", O. Harrassowitz, Wiesbaden. Cité dans Versteegh (2001).
- *Owens, J., 1988, "*The Foundations of Grammar: An Introduction to Medieval Arabic Grammatical Theory*", Benjamins, Amsterdam. Cité dans Versteegh (2001).
- *Owens, J., 1993, "A Grammar of Nigerian Arabic", in *Sprache und Geschichte in Afrika (SGA)*, Vol. 14, pp. 85-175. Cité dans Versteegh (2001).
- *Palva, H., 1969_a, "Balgâwi Arabic 1- Texts from Madaba", in *Studia Orientalia (Helsinki)*, Vol. 40:1, pp. 3-13. Cité dans Barkat (2000).
- *Palva, H., 1969_b, "Balgâwi Arabic 2- Texts in the Dialect of the yigûl-Group", in *Studia Orientalia (Helsinki)*, Vol. 40:2, pp. 3-15. Cité dans Barkat (2000).
- *Palva, H., 1976, "*Studies in the Arabic dialect of the semi-nomadic al-Adjârma tribe (al-Balqâ District, Jordan)*", Acta Universitatis Gothoburgensis, Stockholm. Cité dans Sakarna (2005).
- *Palva, H., 1978, "*Narratives and poems from H'esbân. Arabic texts recorded among the semi-nomadic el-Adjârma tribe (al-Balqâ district, Jordan)*", Acta Universitatis Gothoburgensis, Stockholm. Cité dans Barkat (2000).
- *Palva, H., 1980, "Characteristics of the Arabic Dialect of the Bani Saxar Tribe", in *Orientalia Suecana (OrS)*, Vol. 29, pp. 112-137. Cité dans Sakarna (2005).
- *Palva, H., 1986, "Characteristics of the Arabic Dialect of Hwetat Tribe", in *Orientalia Suecana (OrS)*, Vol. 33-35, pp. 295-312. Cité dans Sakarna (2005).
- *Palva, H., 1989, "Linguistic sketch of the Arabic dialect of el-Karak", in *Studia Linguistica et orientalia Memoriae haim Blanc Dedicata*, Wiesbaden, pp. 225-251. Cité dans Barkat (2000).
- *Palva, H., 1991, "Is there a North West Arabian dialect group?" in *Festgabe für Hans-Rudolf Singer*, M. F. (Ed), P. Lang, Frankfurt am main, pp. 151-166. Cité dans Versteegh (2001).
- *Palva, H., 1993, "Bedouin and sedentary elements in the dialect of es-Salt' - Diachronic notes on the sociolinguistic development", in *Actes des premières journées internationales de dialectologie arabe de Paris*, INALCO (Paris), Langues'O, pp. 459-469. Cité dans Barkat (2000).
- *Peterson, G. E., 1961, "Parameters of vowel quality", in *Journal of Speech and Hearing Research*, Vol. 4, pp. 10-29. Cité dans Strange (1989).
- *Rabin, C., 1951, "*Ancient West Arabian*", Taylor's Foreign Press, London. Cité dans Sakarna (2005).
- *Rabin, C., 1955, "The beginnings of Classical Arabic", in *Studia Islamica*, Paris, Vol. 4, pp. 19-37. Cité dans Versteegh (2001).
- *Roman, A., 1983, "*Étude de la phonologie et de la morphologie de la Koinè arabe : tomes 1 et 2*", Publications de l'Université de Provence, Aix-en-Provence, France. Cité dans Versteegh (2001).

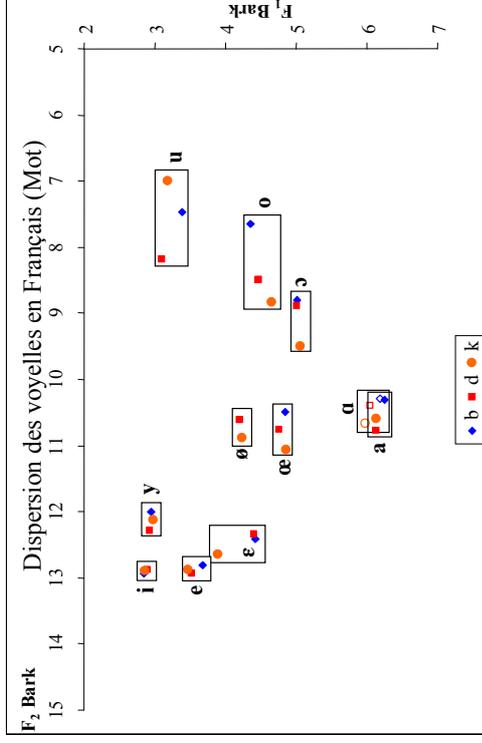
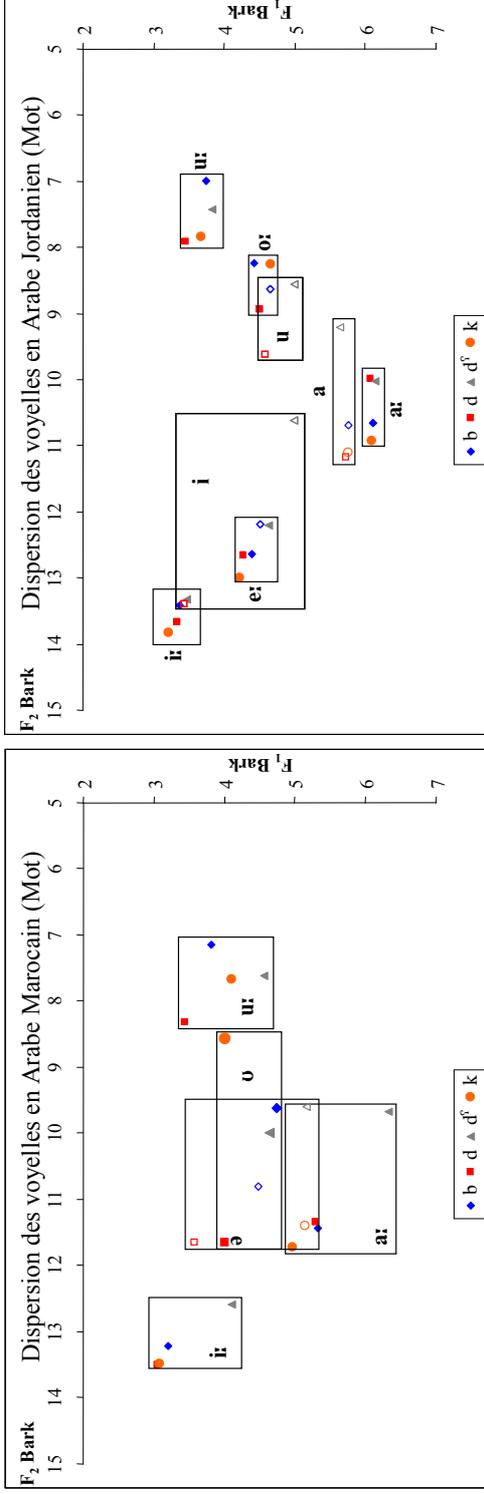
- *Rosenhouse, J., 1984, "*The Bedouin Arabic Dialects - General Problems and a close analysis of North Israel Bedouin Dialects*", Otto Harrassowitz Publishers, Wiesbaden. Cité dans Sakarna (2005).
- *Sakarna, A. K., 1997, "*The Linguistic Status of the Modern Jordanian Dialects*", University of Wisconsin-Madison. Cité dans Sakarna (1999).
- *Selfridge, O. G., 1959, "Pandemonium: A paradigm for learning", In *Proc. of the NPL Symp. 1958, No. 10, Mechanization of Thought Processes (Her Majesty's Stationery Office, London, 1959)*, pp. 513-526. Cité dans Sanders (1977).
- *Singer, H.-R., 1984, "*Grammatik der arabischen Mundart der Medina von Tunis*", De Gruyter, Berlin & New York. Cité dans Versteegh (2001).
- *Singer, H.-R., 1994, "Die Beduinen als Träger der Arabisierung im islamischen Machtbereich", in *Gedenkschrift Wolfgang Reuschel. Akten des III. Arabistischen Kolloquiums*, Leipzig, 21-22 November 1991, Dieter Bellmann (Eds), pp. 263-274, F. Steiner, Stuttgart. Cité dans Versteegh (2001).
- *Stevens, K. N., and Halle, M., 1967, "Remarks on analysis by synthesis and distinctive features", in *Models for the perception of speech and visual form*, W. Wathen-Dunn (Ed), MIT Press, Cambridge, MA, pp. 88-102. Cité dans Sanders (1977).
- *Stevens, K. N., and House, A. S., 1963, "Perturbation of vowel articulations by consonantal context: An acoustical study", in *Journal of Speech and Hearing Research*, Vol. 6, pp. 111-128. Cité dans Hillenbrand *et al.* (2001).
- *Stevens, K. N., and House, A. S., 1972, "Speech perception", in *Foundations of modern auditory theory, Vol. II*, J. V. Tobias (Ed), Academic Press, New York, pp. 3-62. Cité dans Sanders (1977).
- *Strange, W., Edman, T. R., and Jenkins, J. J., 1979, "Acoustic and phonological factors in vowel identification", in *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 5, pp. 643-656. Cité dans Strange (1999_a).
- *Suleiman, S. M., 1985, "Jordanian Arabic between diglossia and bilingualism: Linguistic analysis", in *Pragmatics & Beyond*, Vol. 6:8, pp. 1-131. Cité dans Versteegh (2001).
- *Taine-Cheikh, C., 1978, "*L'arabe médian parlé par les arabophones de Mauritanie - Étude morphosyntaxique*", Thèse de 3^{ème} cycle Paris V - René Descartes, Paris. Cité dans Barkat (2000).
- *Taine-Cheikh, C., 1991, "L'arabe des *Bidân*, un dialecte bédouin de Maghreb occidental", in *Semitic Studies (Hommage à Leslau)*, A. S. Kaye (Ed), Otto Harrassowitz, Wiesbaden, pp. 1528-1548. Cité dans Barkat (2000).
- *Talmoudi, F., 1980, "*The Arabic Dialect of Susa (Tunisia)*", Acta Universitatis Gothoburgensis, Göteborg. Cité dans Versteegh (2001).
- *Talmoudi, F., 1981, "*Texts in the Arabic Dialect of Susa (Tunisia)*", Acta Univ. Gothoburgensis, Gothenburg. Cité dans Versteegh (2001).
- *Talmoudi, F., 1984, "*The Diglossic Situation in North Africa: A Study of Classical Arabic/Dialectal Arabic Diglossia with Sample Text in 'Mixed Arabic'*", Acta Univ. Gothoburgensis, Gothenburg. Cité dans Versteegh (2001).
- *Uttley, A. M., 1959, "Conditional Probabilities Computing in a Nervous System", in *Mechanisation of Thought Processes*, Her Majesty's Stationery Office, London. Cité dans Sanders (1977).
- *Uttley, A. M., 1966, "The transmission of information and the effect of local feedback in theoretical and neural networks", in *Brain Research*, Vol. 2, pp. 21-30. Cité dans Sanders (1977).

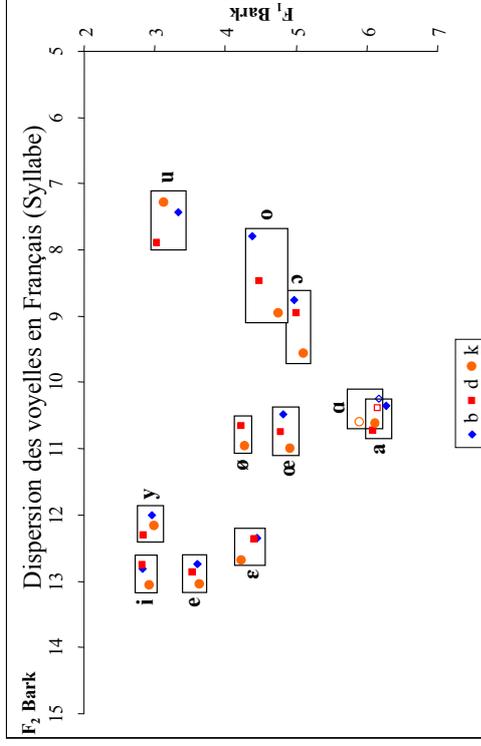
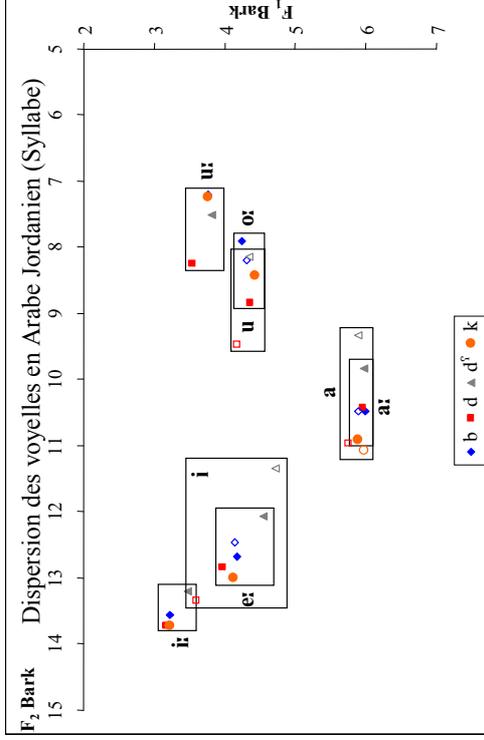
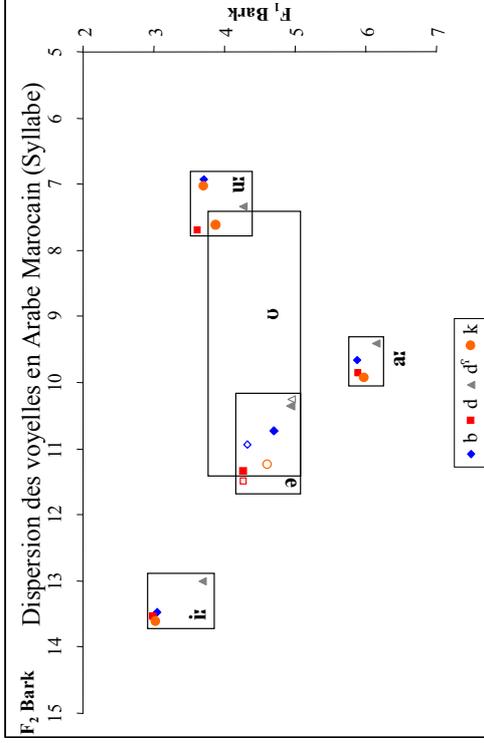
- *van Bezooijen, R., and van den Berg, R., 1999, "Word intelligibility of language varieties in the Netherlands and Flanders under minimal conditions", in *Linguistics in the Netherlands*, Amsterdam: John Benjamins, Vol. 16, pp. 1-12. Cité dans Thomas (2002).
- *Vanhove, M., 1995_a, "A propos du verbe dans les dialectes arabes de YAfi' (Yémen)" in *Dialectologia Arabica. A Collection of Articles in Honour of the Sixtieth Birthday of Professor Heikki Palva*, pp. 257-269. Cité dans Vanhove (2002).
- *Vanhove, M., 1995_b, "Notes on the Arabic Dialectal Area of YAfi' (Yemen) ", in *Proc. of the Seminar for Arabian Studies (Oxford, July 1994)*, Vol. 25, pp. 141-152. Cité dans Vanhove (2002).
- *Vanhove, M., 1996_a, "The negation maa'ii in a Yaafifi dialect (Yemen)", in *Perspectives on Arabic Linguistics*, Vol. 9, pp. Mushira Eid, and D. Parkinson (Eds), Amsterdam/Philadelpia, John Benjamins, pp. 195-206. Cité dans Vanhove (2002).
- *Vanhove, M., 1997, "Diversité dialectale et nivellement linguistique : le cas de l'arabe au Yémen (sud et est)", in *Proc. of the 16th International Congress of Linguists*, Paris 20-25 July 1997, B. Caron (Ed), Pergamon, Oxford, Paper n° 115. Cité dans Vanhove (2002).
- *Verbrugge, R. R., and Rakerd, B., 1986, "Evidence of talker-independent information for vowels", in *Language and Speech*, Vol. 29, pp. 39-57. Cité dans (Strange (1989).
- *Versteegh, K., 1985, "The Development of Argumentation in Arabic Grammar: The Declension of the Dual and the Plural", in *Zeitschrift für Arabische Linguistik (Z.A.L)*, Vol. 15, pp. 152-173. Cité dans Versteegh (2001).
- *Versteegh, K., 1995, "*The Explanation of Linguistic Causes: Az-Zağğāğī's Theory of Grammar, Introduction, Translation, and Commentary*", J. Benjamins, Amsterdam and Philadelphia. Cité dans Versteegh (2001).
- *Wickelgren, W. A., 1969, "Context-sensitive coding, associative memory, and serial order in (speech) behavior", in *Psychological Review*, Vol. 76, pp. 1-15. Cité dans Tatham & Morton (2006).
- *Woidich, M., 1979, "Zum Dialekt von il-'Awâmra in der östlichen sharqiyya (Ägypten). Vol. I : Einleitung, grammatische Skizzen und volkskundliches", in *Zeitschrift für Arabische Linguistik (Z.A.L)*, Vol. 2, pp. 76-99. Cité dans Versteegh (2001).
- *Woidich, M., 1980, "Zum Dialekt von il-'Awâmra in der östlichen sharqiyya (Ägypten). Vol. II : Texte und Glossar", in *Zeitschrift für Arabische Linguistik (Z.A.L)*, Vol. 4, pp. 31-60. Cité dans Versteegh (2001).
- *Woidich, M., 1990, "*Ahlan wa sahan: Eine Einführung in die Kairoer Umgangssprache*", L. Reichert, Wiesbaden. Cité dans Versteegh (2001).
- *Woidich, M., 1993, "Die Dialekte der ägyptischen Oasen: westliches oder östliches Arabisch?" in *Zeitschrift für Arabische Linguistik (Z.A.L)*, Vol. 25, pp. 340-359. Cité dans Versteegh (2001).
- *Younes, M., 1982, "*Problems in the Segmental Phonology of Palestinian Arabic*", Ph.D. Dissertation, University of Texas, Austin. Cité dans Sakarna (1999).
- *Youssi, A., 1983, "La triglossie dans la typologie linguistique", in *La Linguistique*, Vol. 19:2, pp. 71-83. Cité dans Ennaji (2002_a).
- *Youssi, A., 1995, "The Morrocan triglossia: Facts and implications", in *International Journal of the Sociology of Language*, Vol. 112: Sociolinguistics in Morroco (special issue), pp. 29-43. Cité dans Ennaji (2002_a).

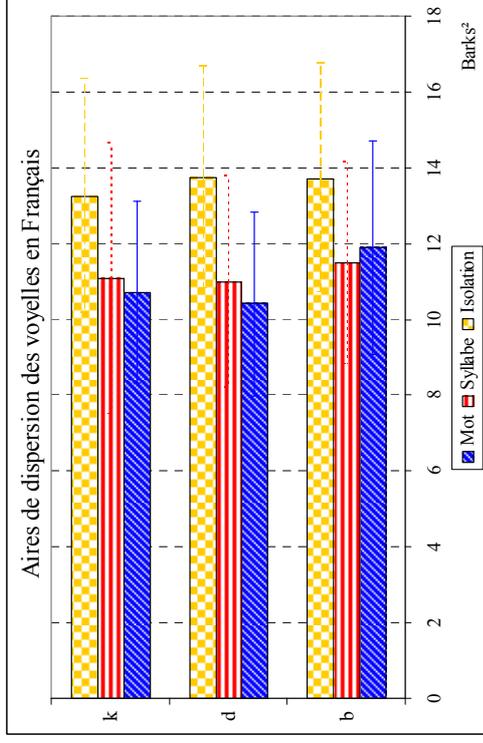
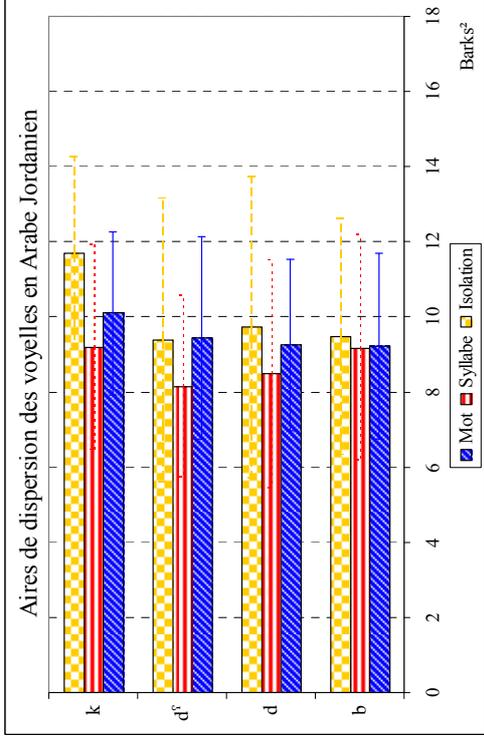
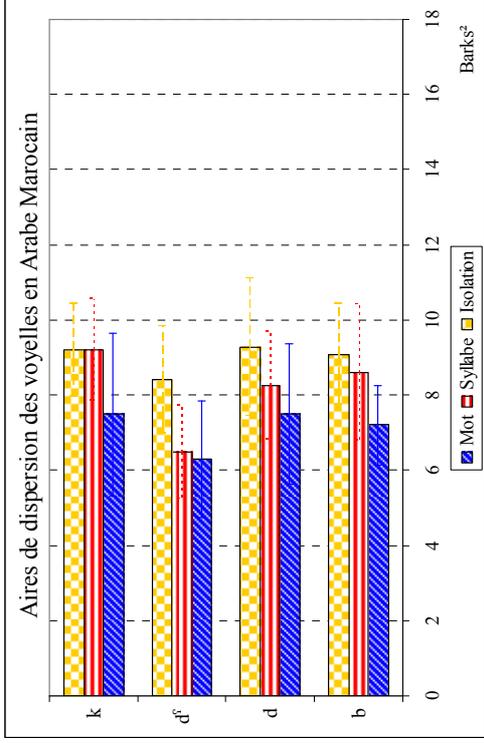
ANNEXES

ANNEXE 1 ESPACES VOCALIQUES ET AIRES DE DISPERSION DES VOYELLES EN ARABE MAROCAIN, ARABE JORDANIEN ET EN FRANÇAIS

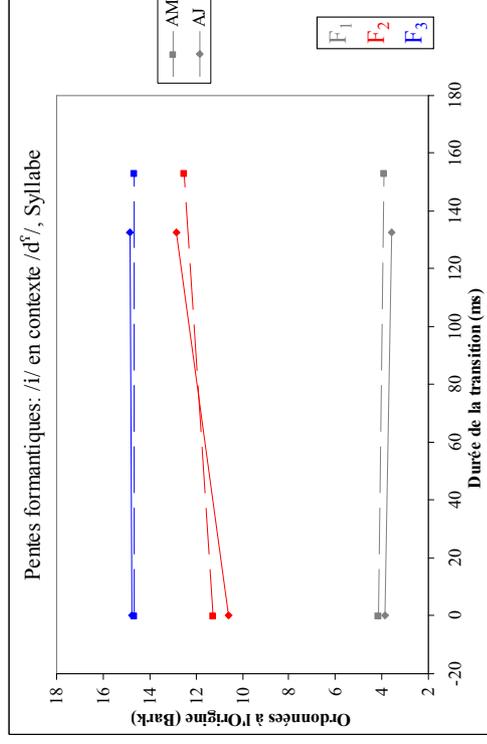
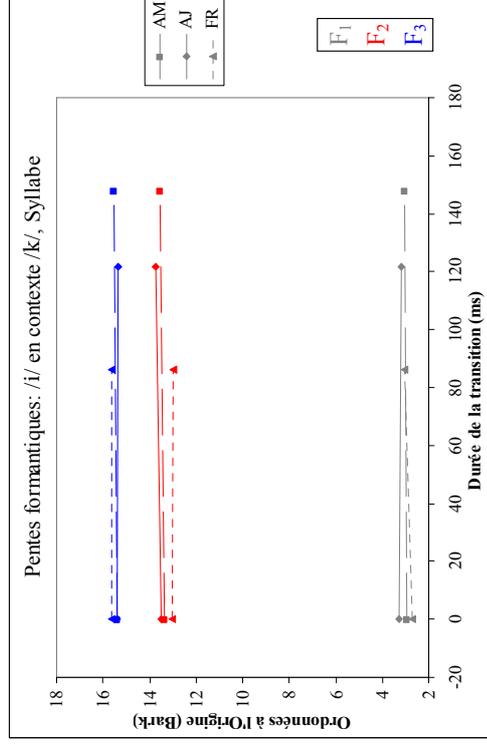
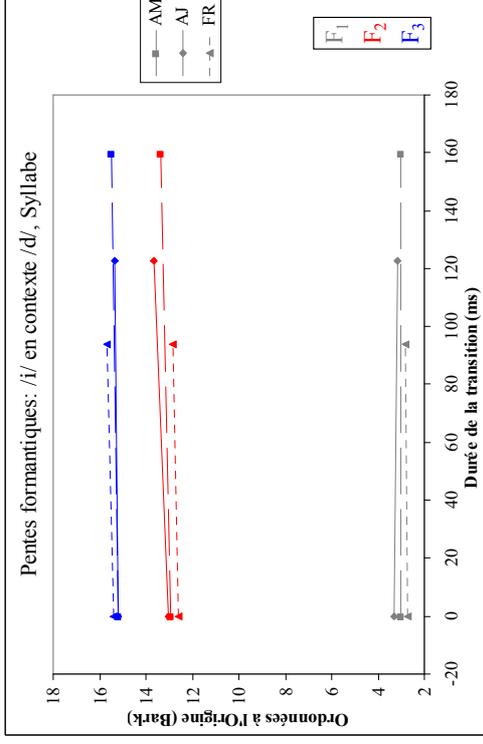
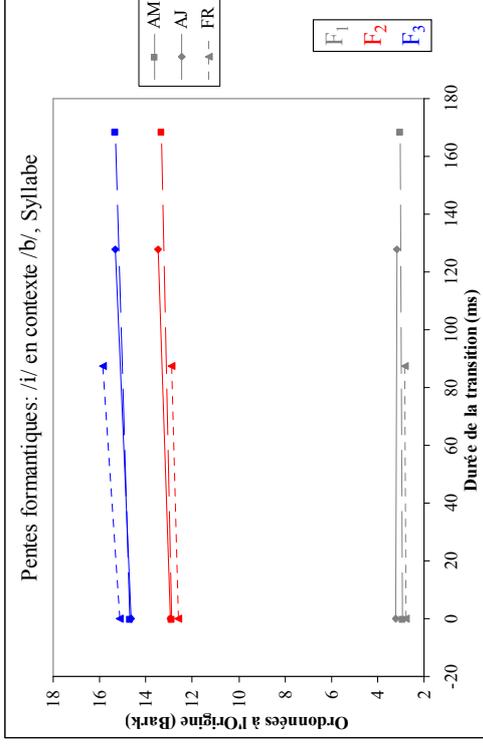


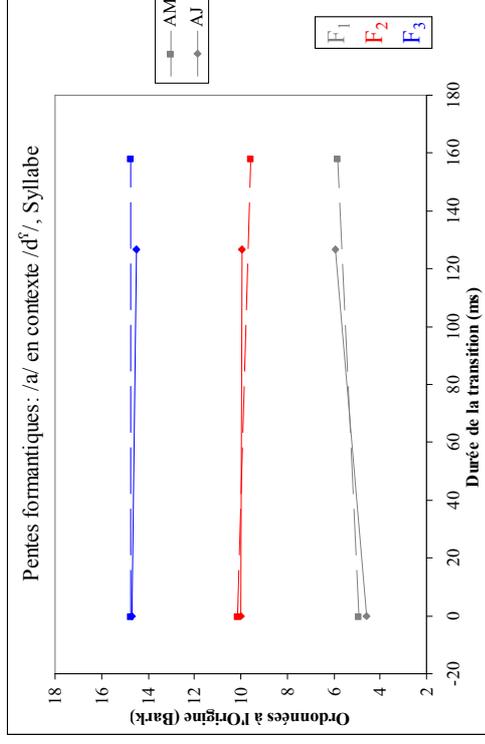
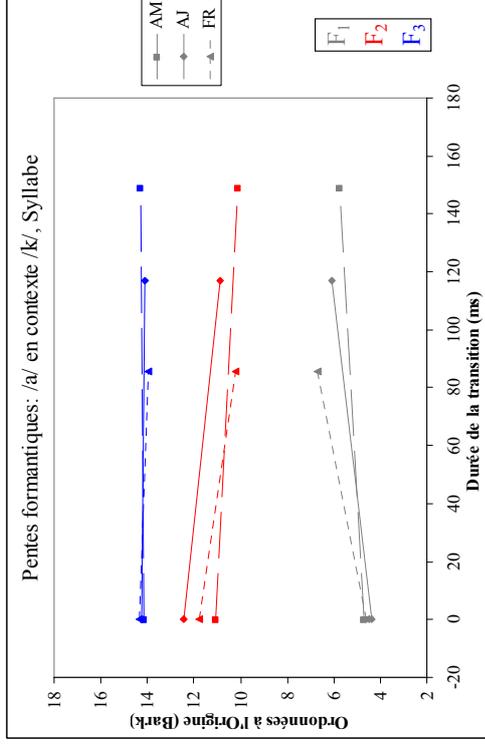
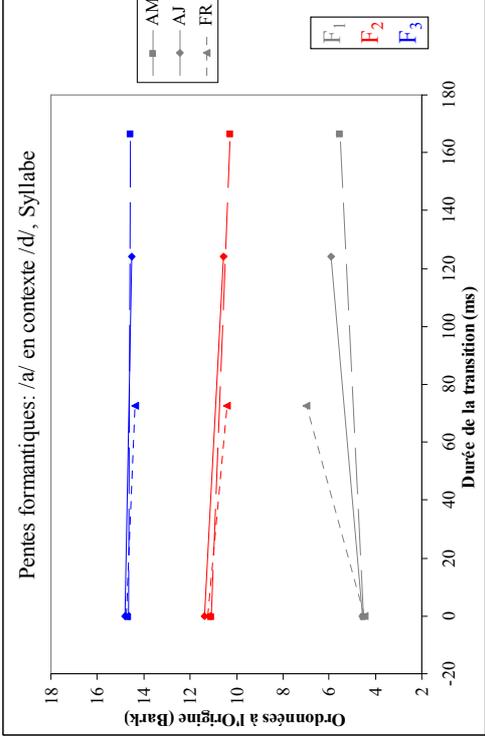
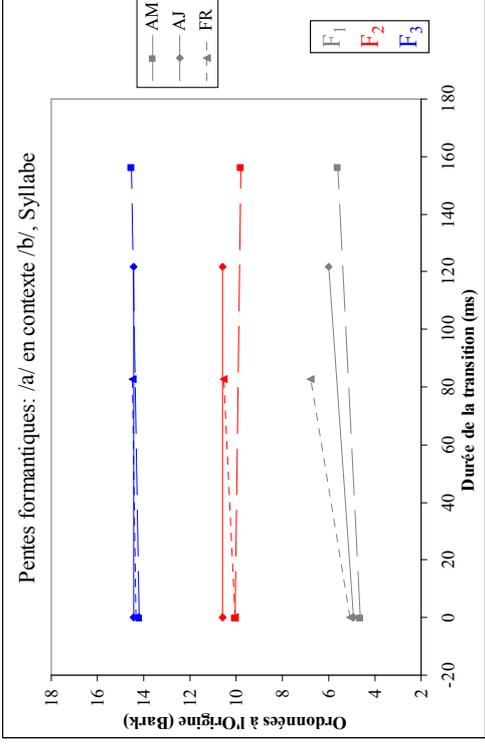


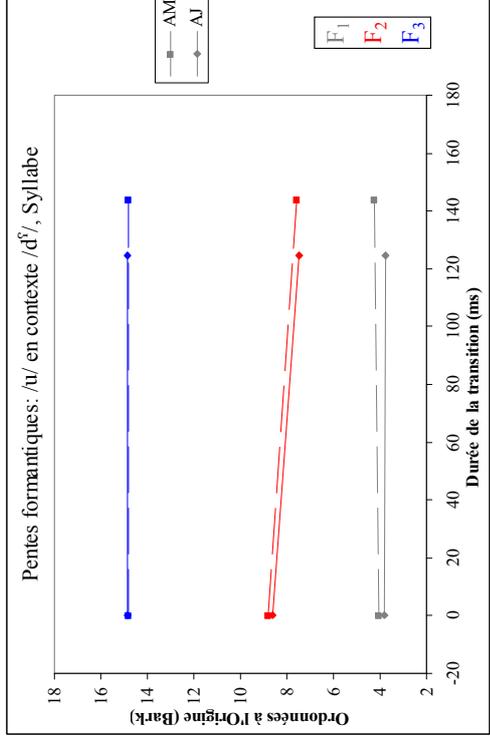
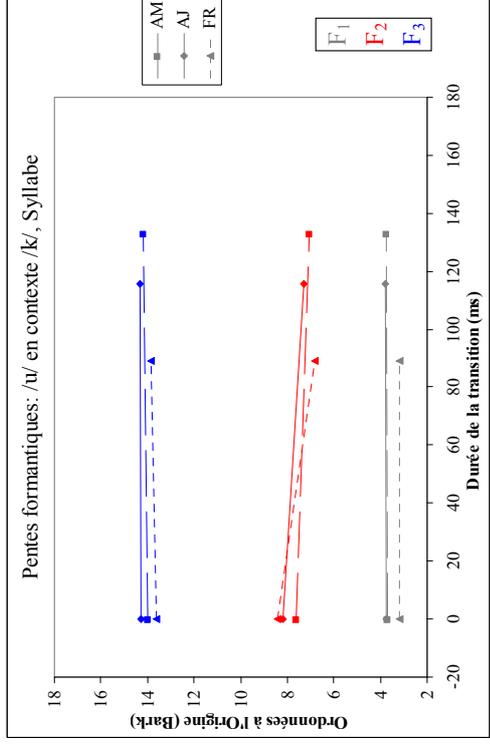
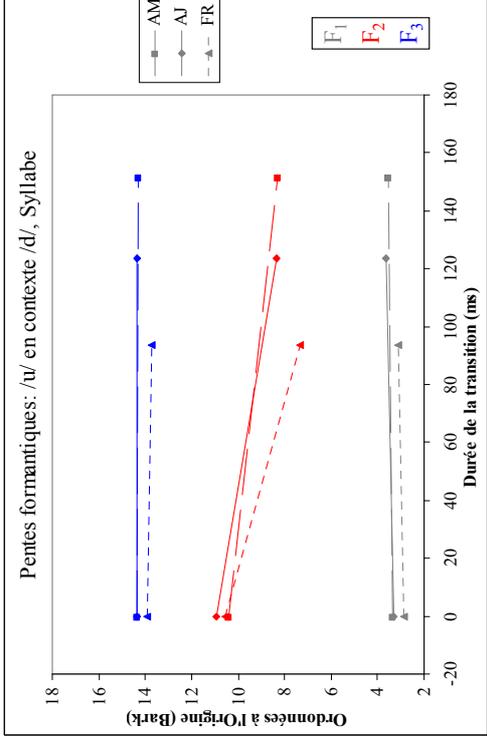
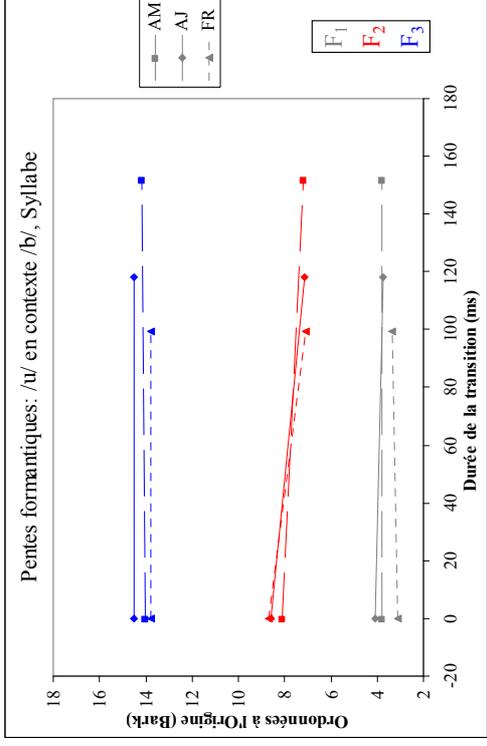




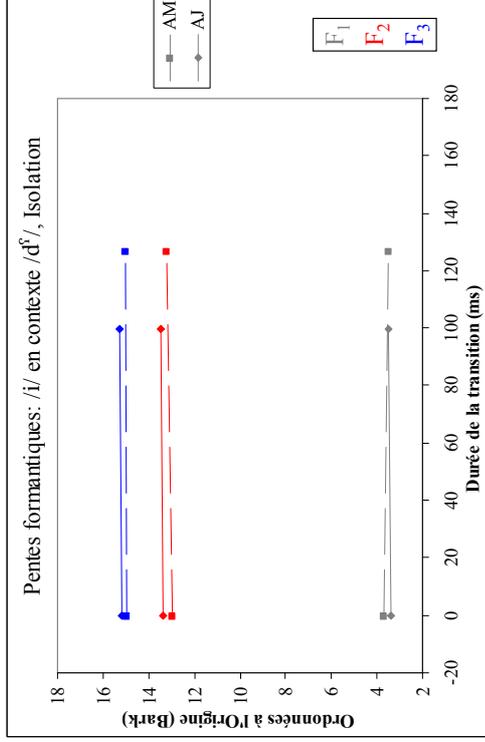
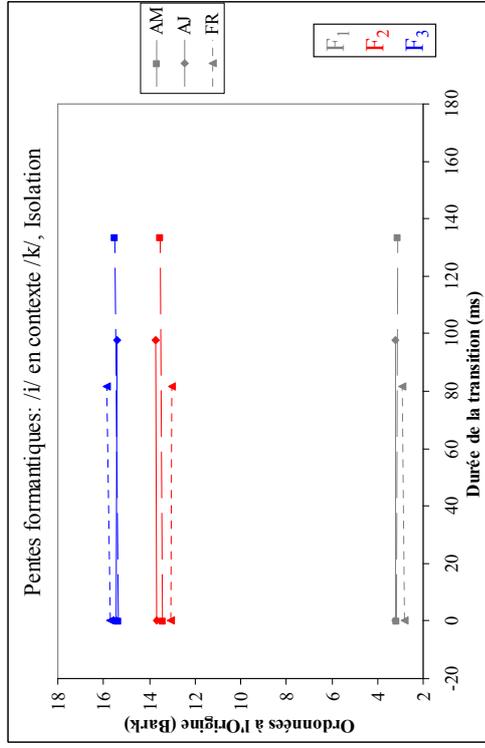
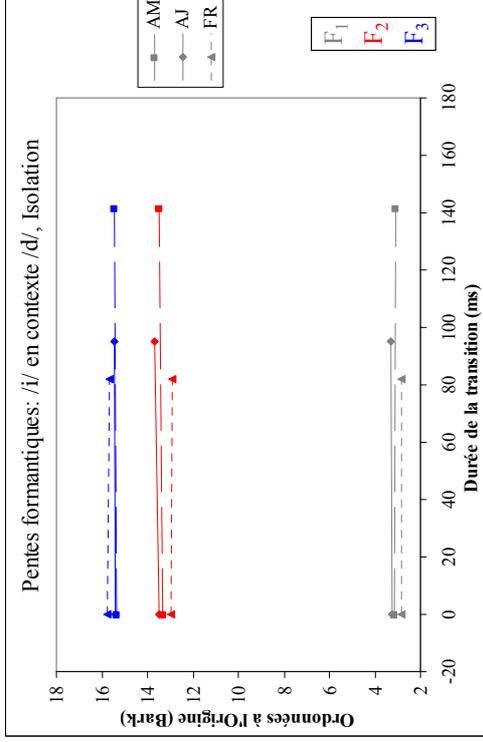
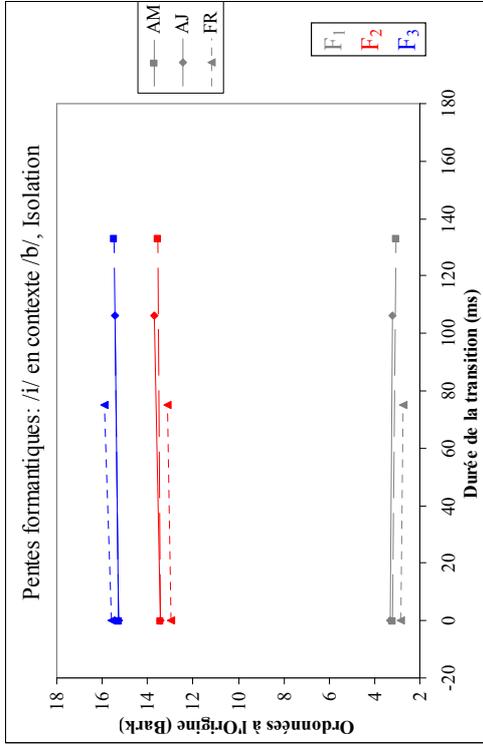
ANNEXE 2 PENTES FORMANTIQUES SANS NORMALISATION DU TEMPS EN SITUATION SYLLABE EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS

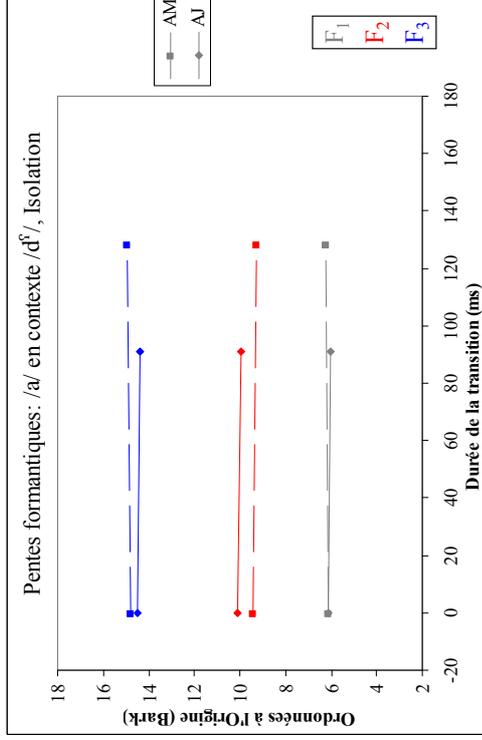
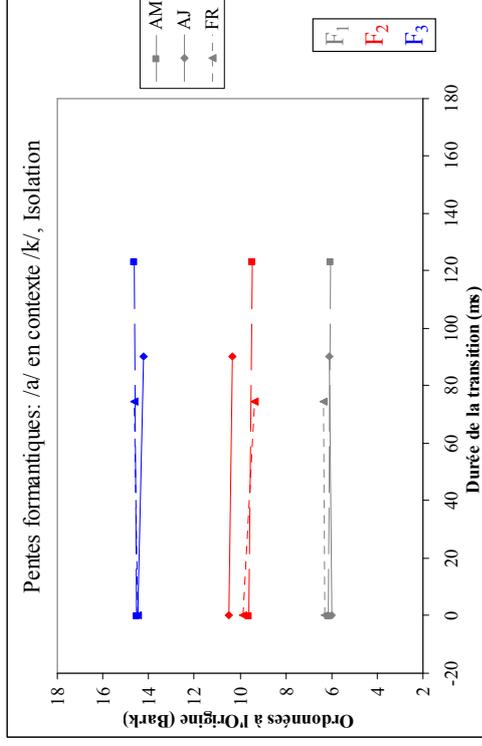
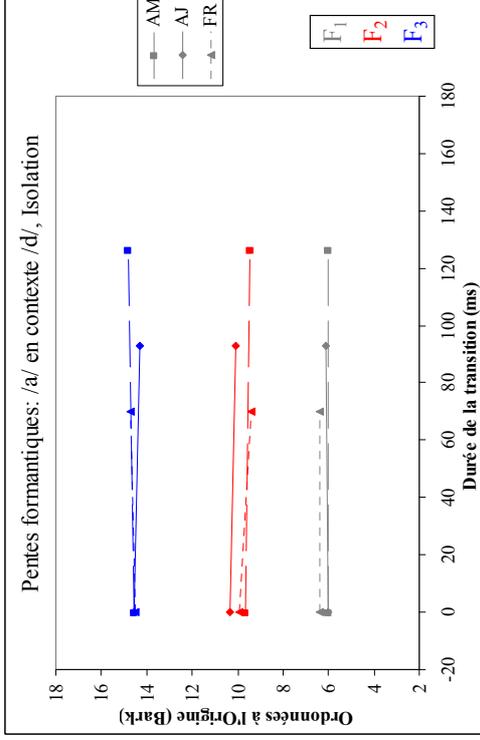
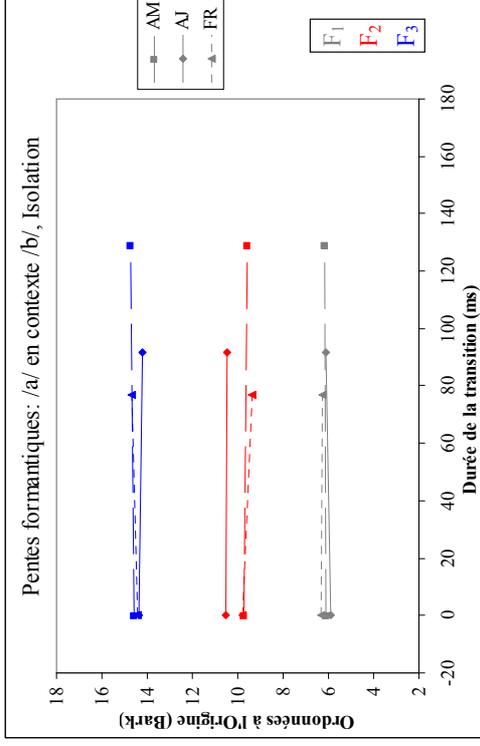


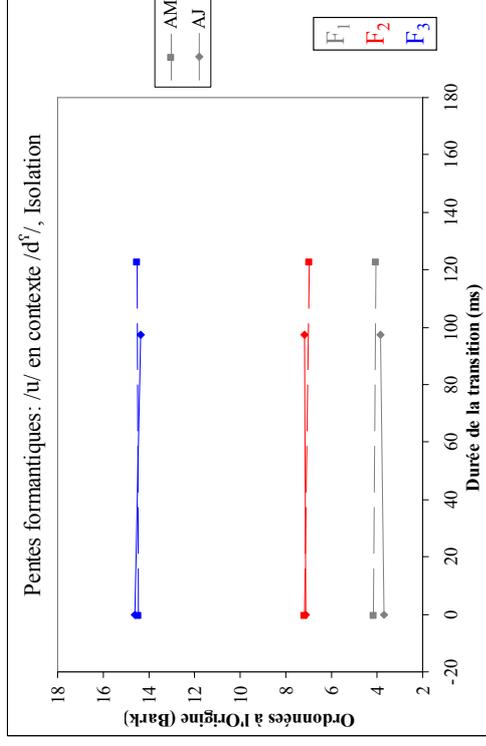
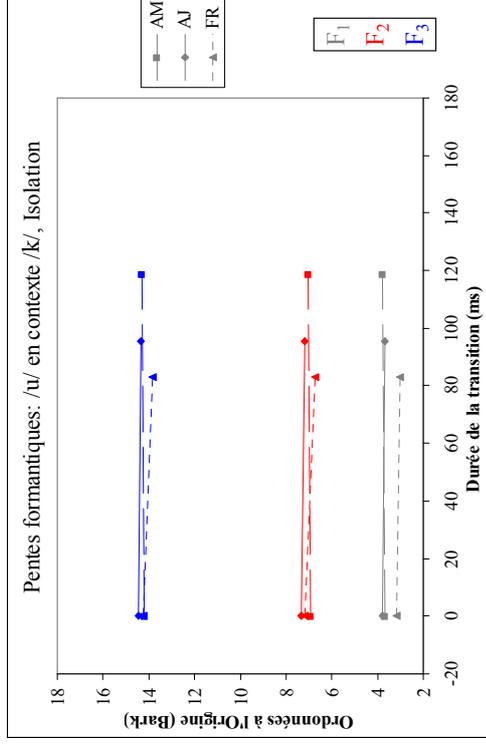
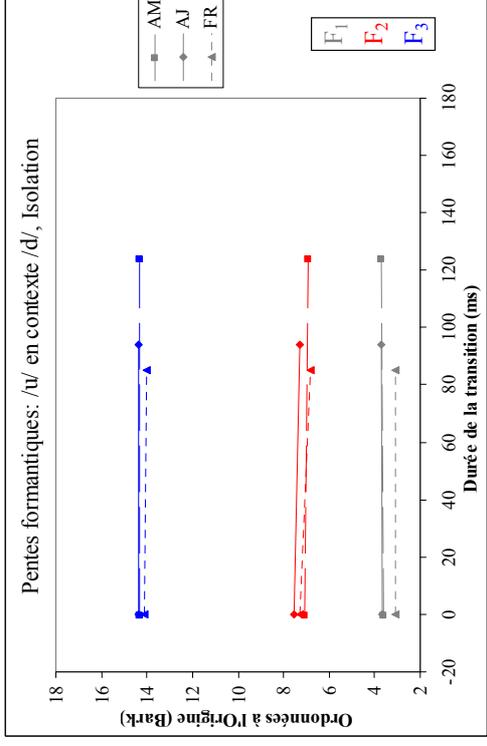
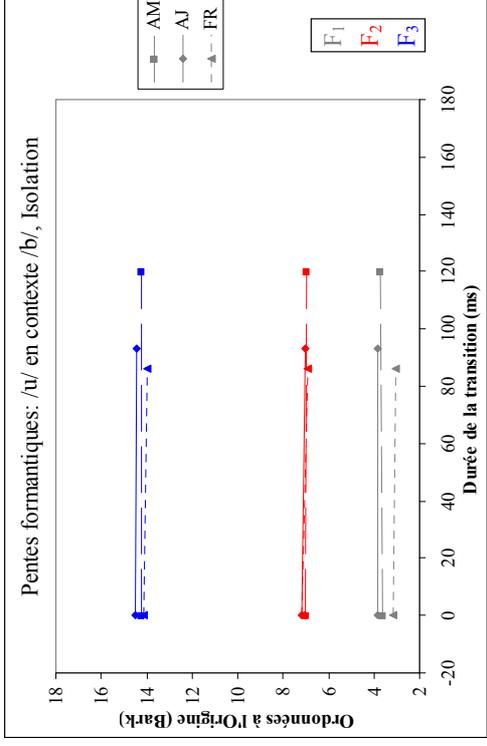




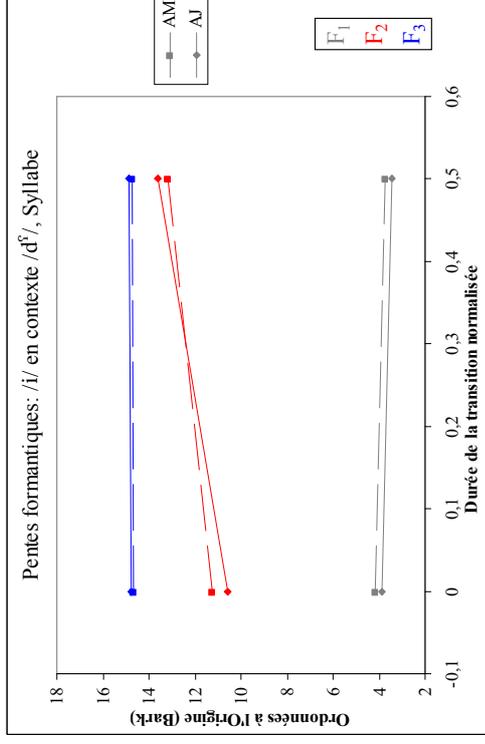
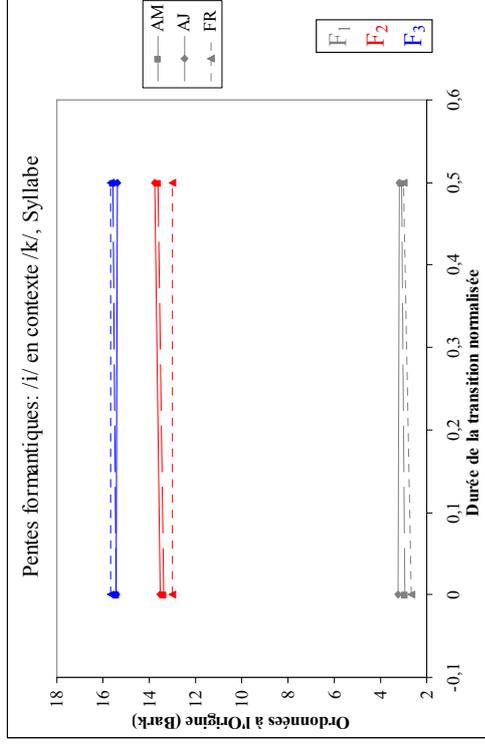
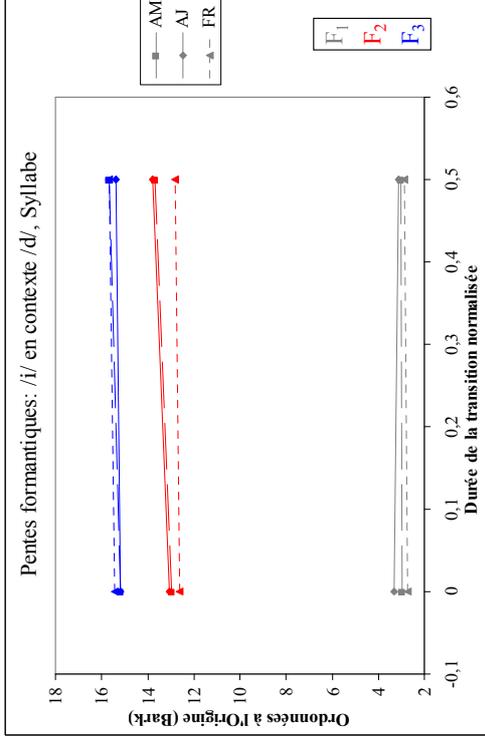
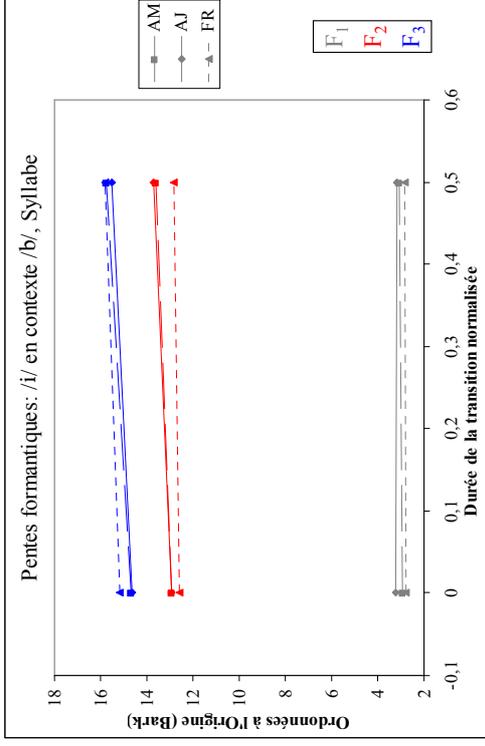
ANNEXE 3 PENTES FORMANTIQUES SANS NORMALISATION DU TEMPS EN SITUATION ISOLATION EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS

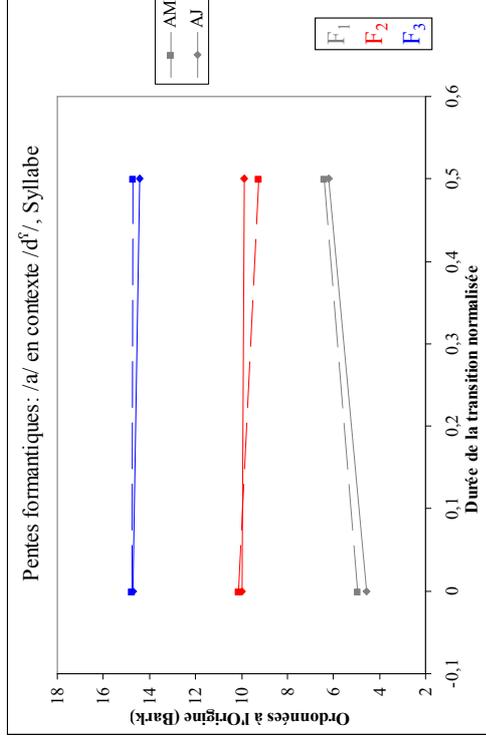
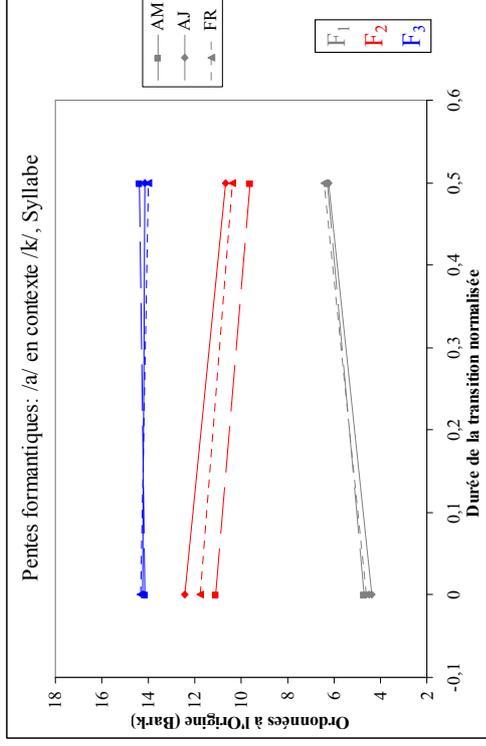
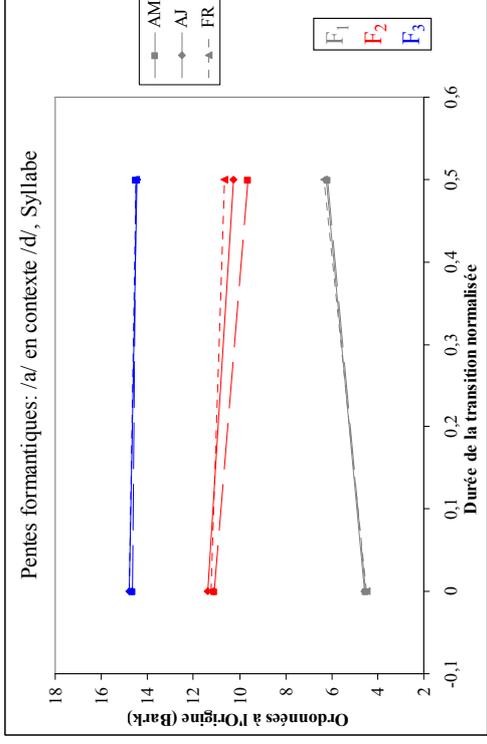
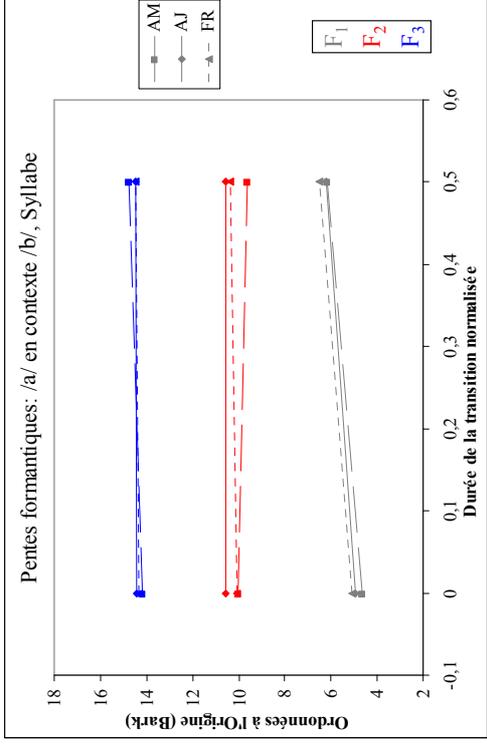


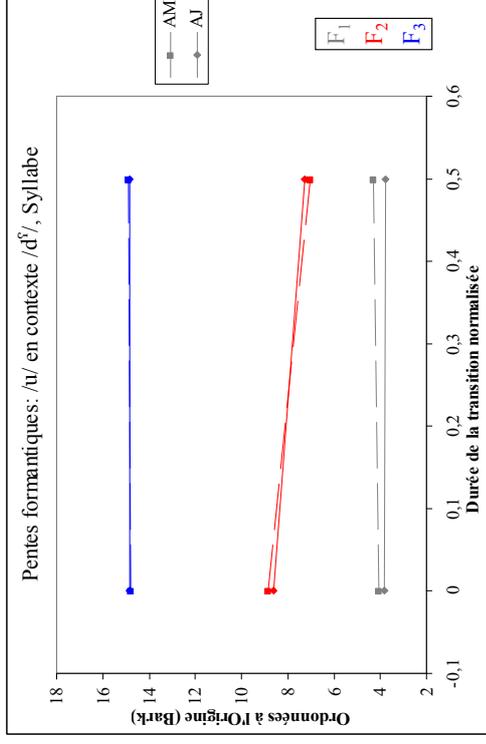
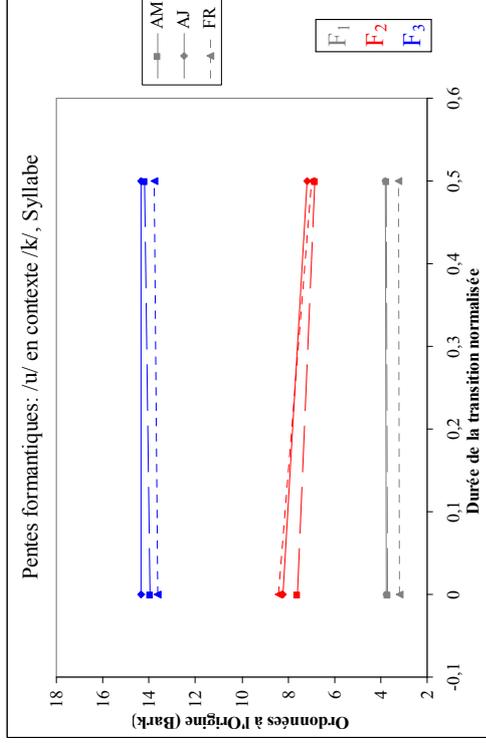
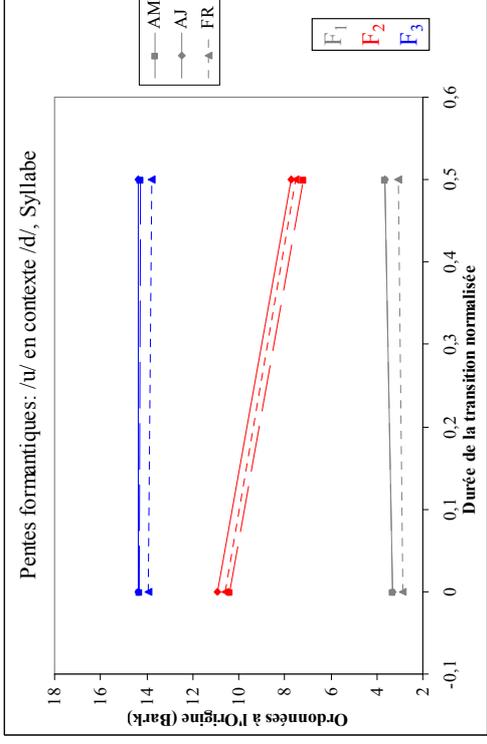
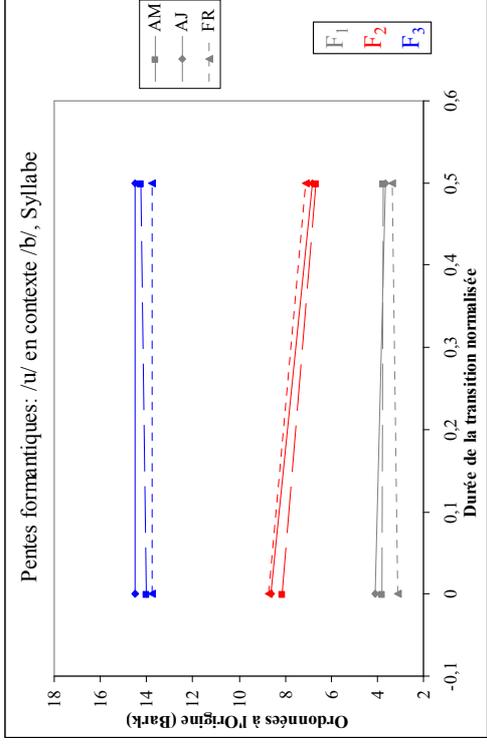




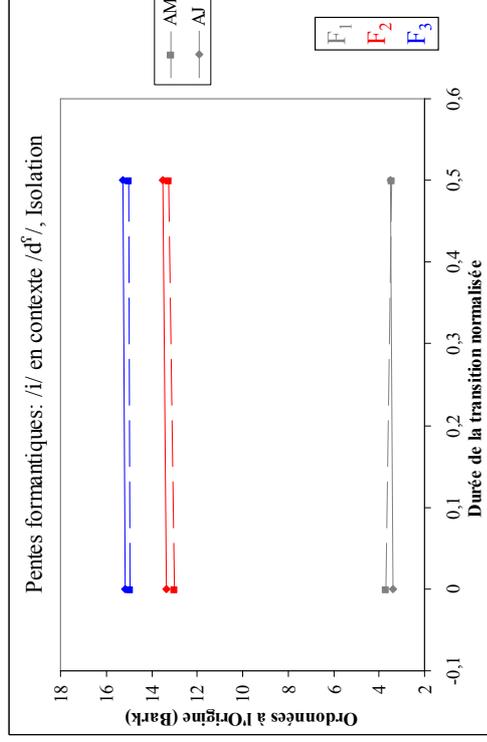
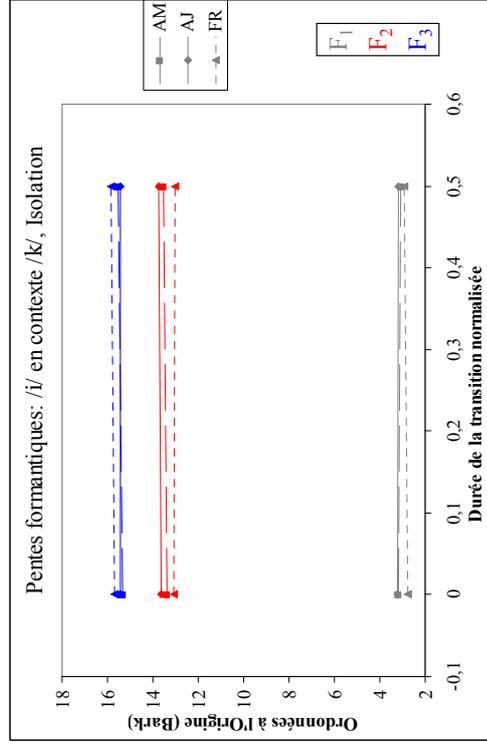
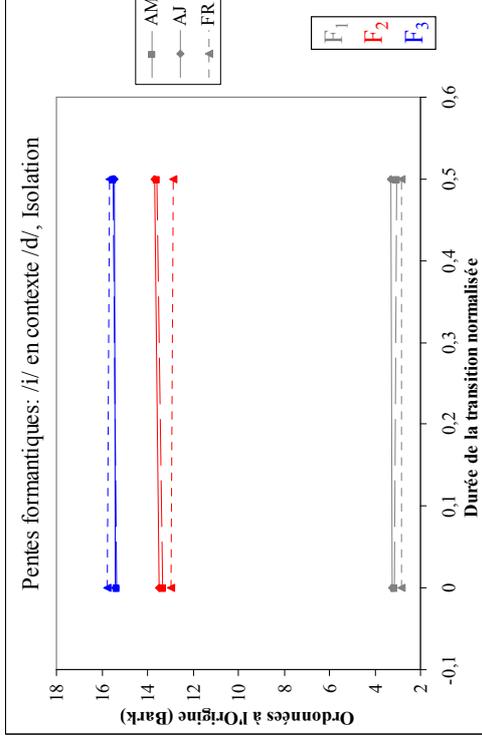
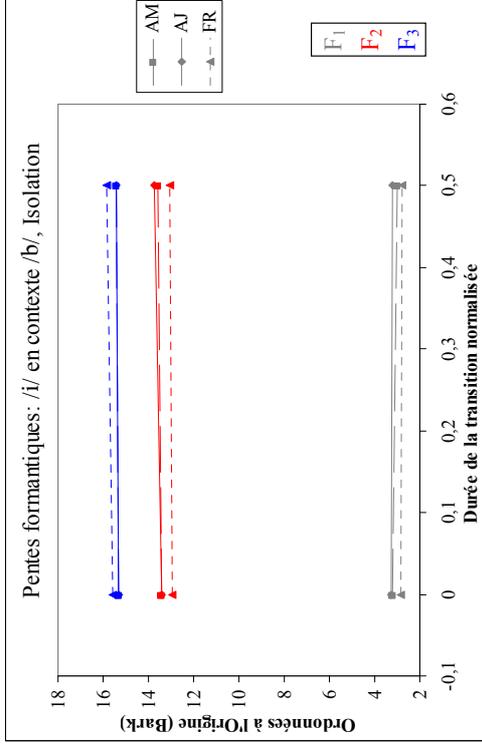
ANNEXE 4 PENTES FORMANTIQUES AVEC NORMALISATION DU TEMPS EN SITUATION SYLLABE EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS

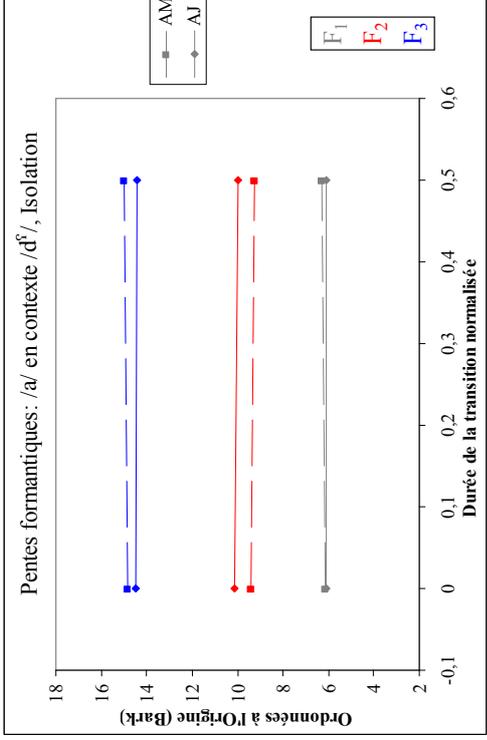
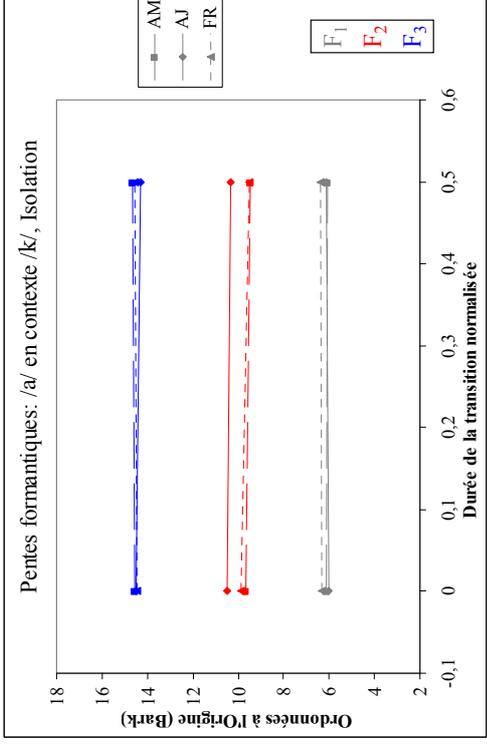
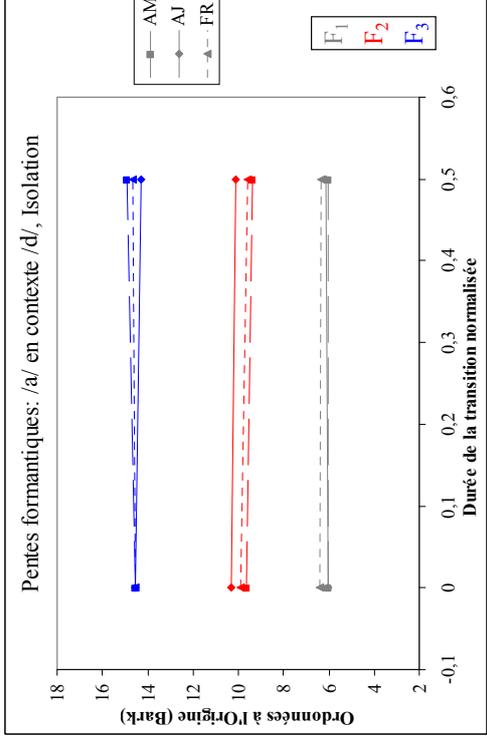
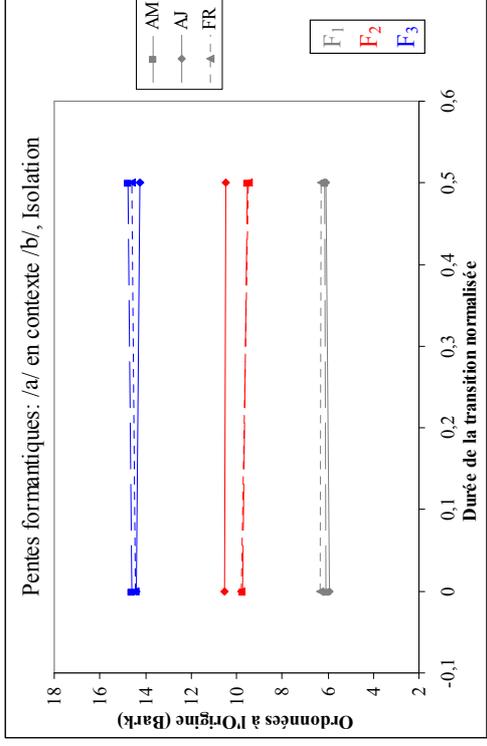


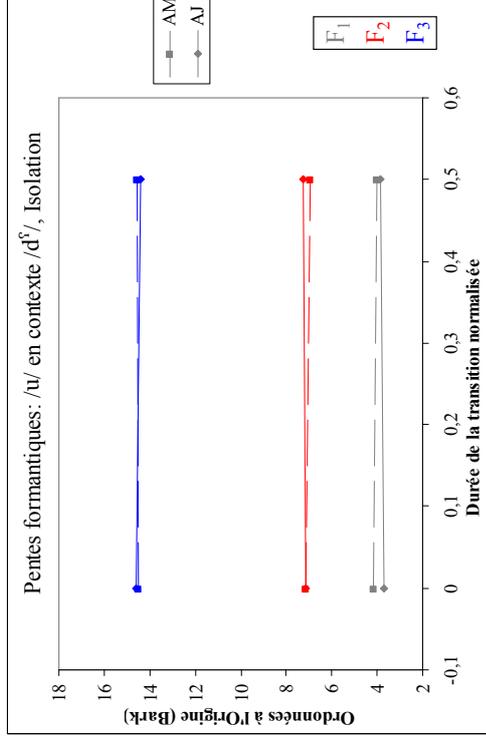
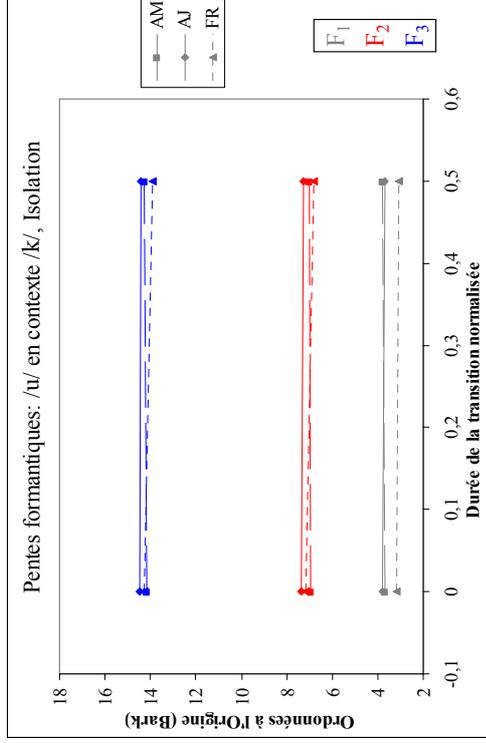
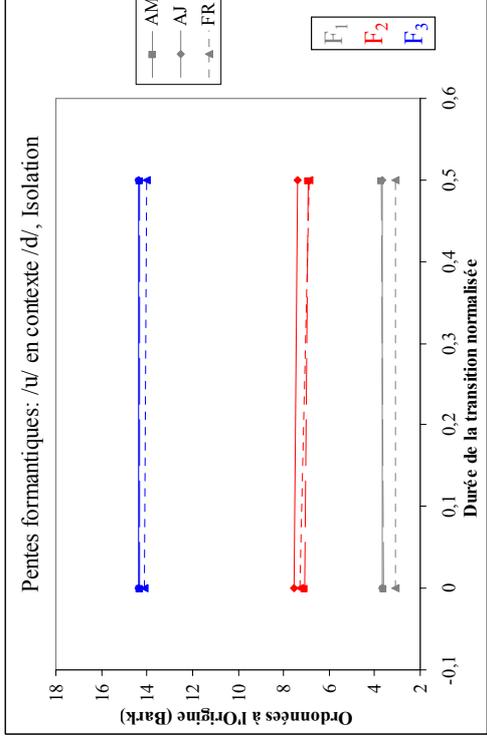
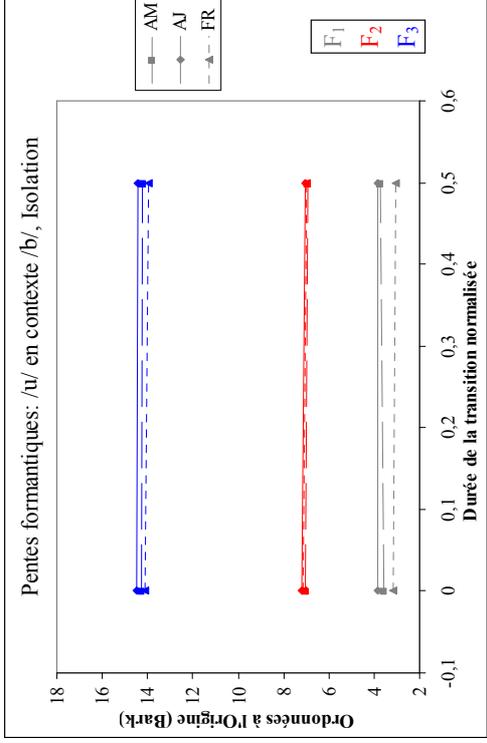




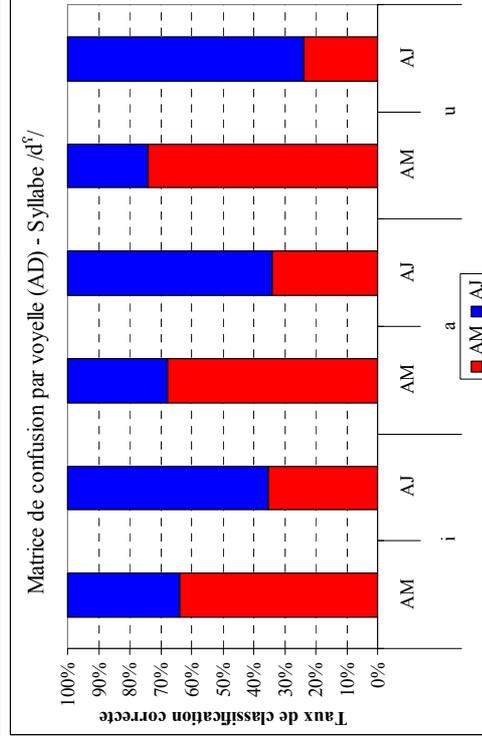
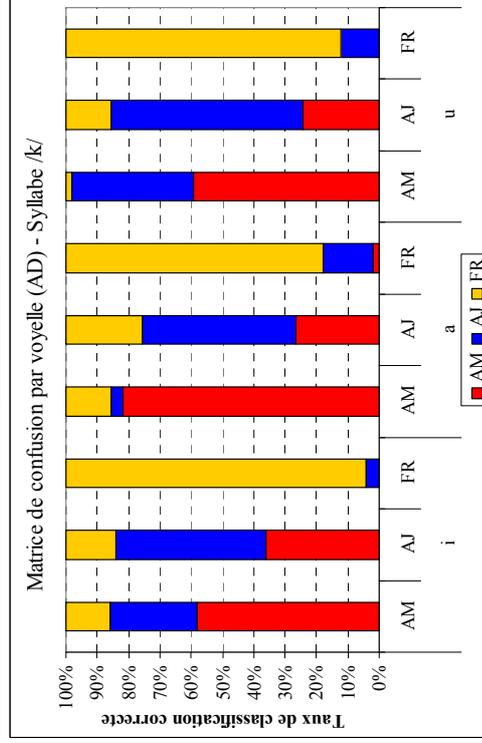
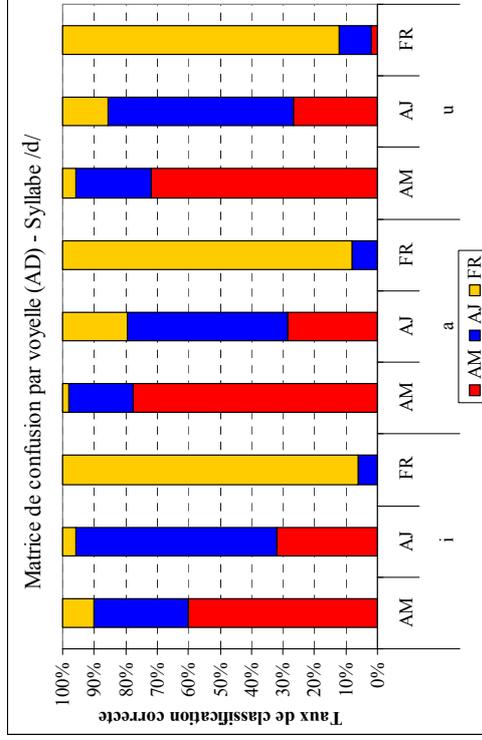
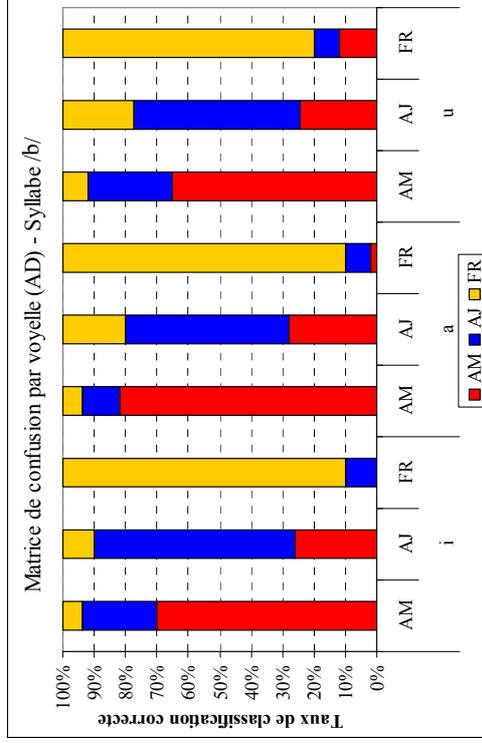
ANNEXE 5 PENTES FORMANTIQUES AVEC NORMALISATION DU TEMPS EN SITUATION ISOLATION EN ARABE MAROCAIN, ARABE JORDANIEN & EN FRANÇAIS

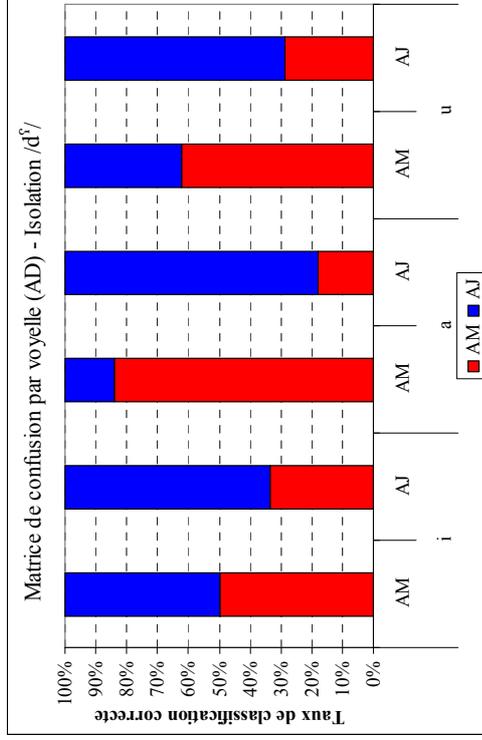
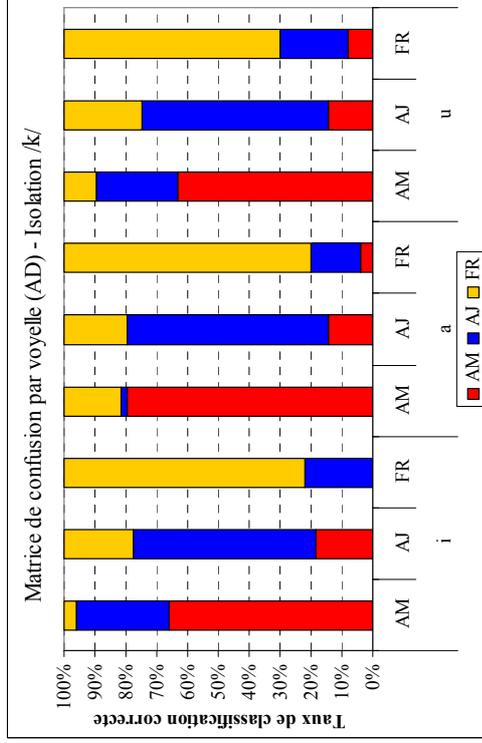
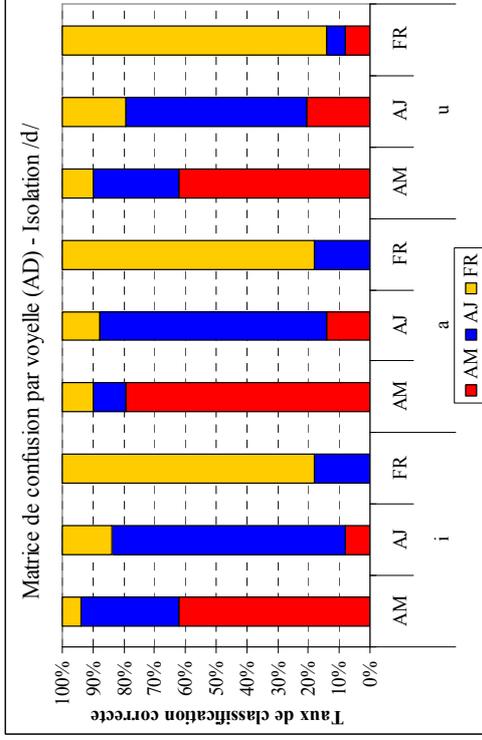
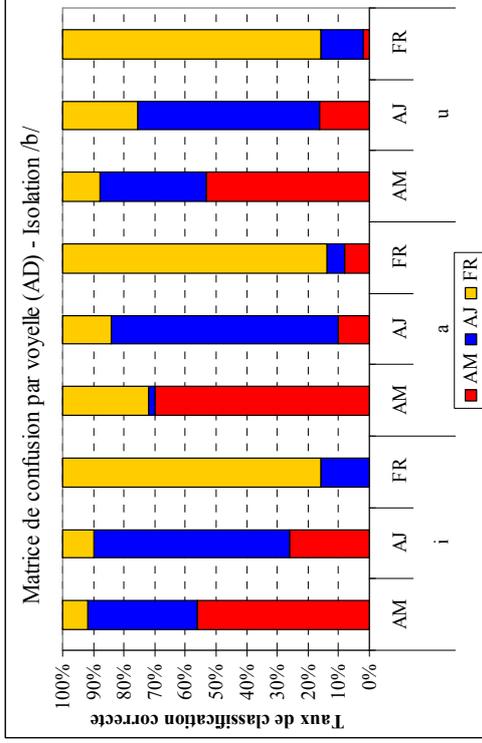




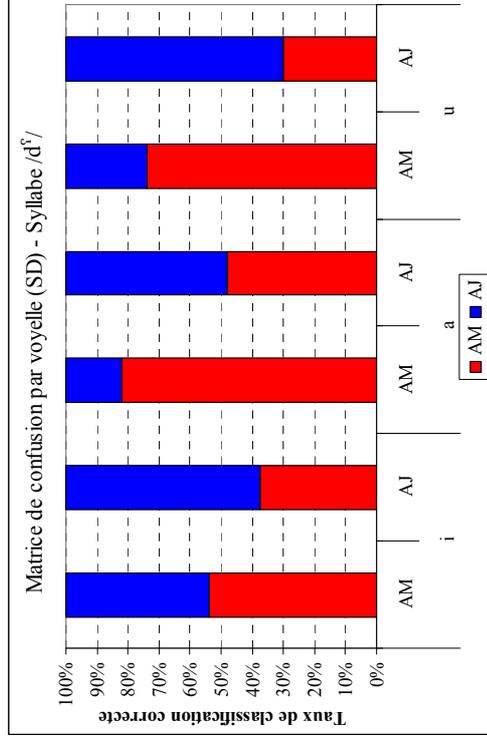
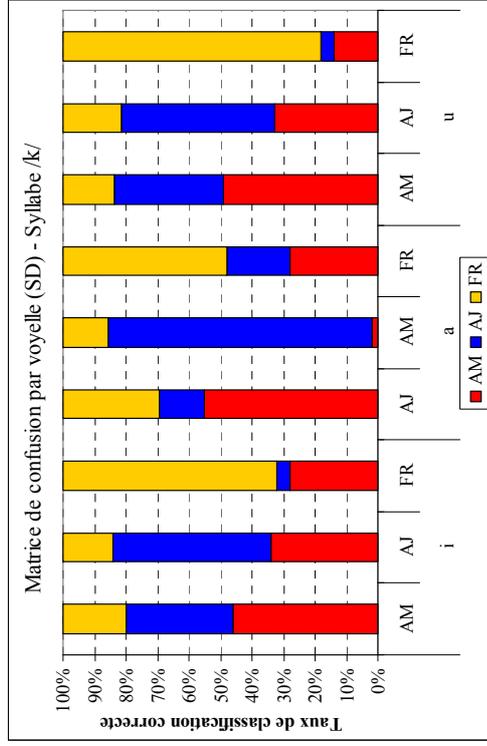
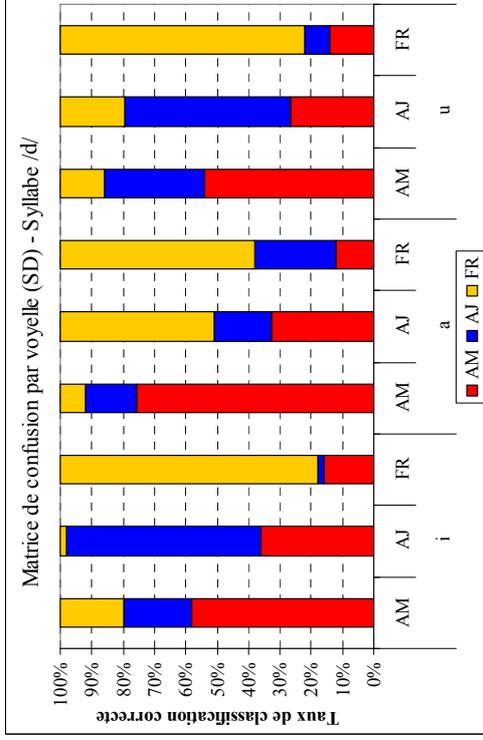
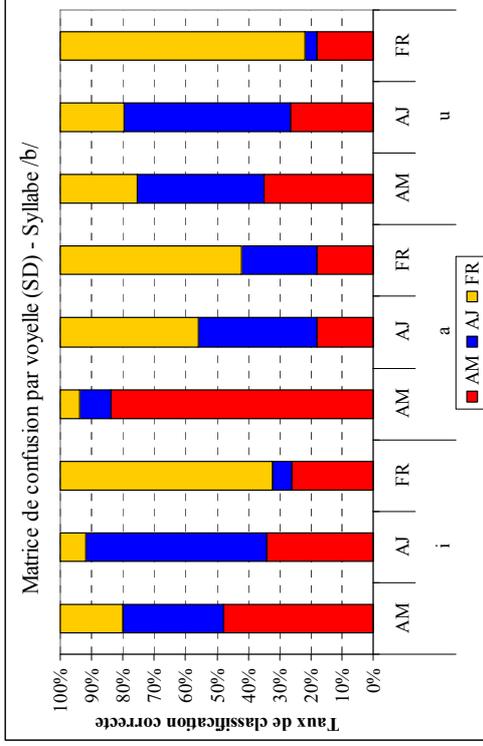


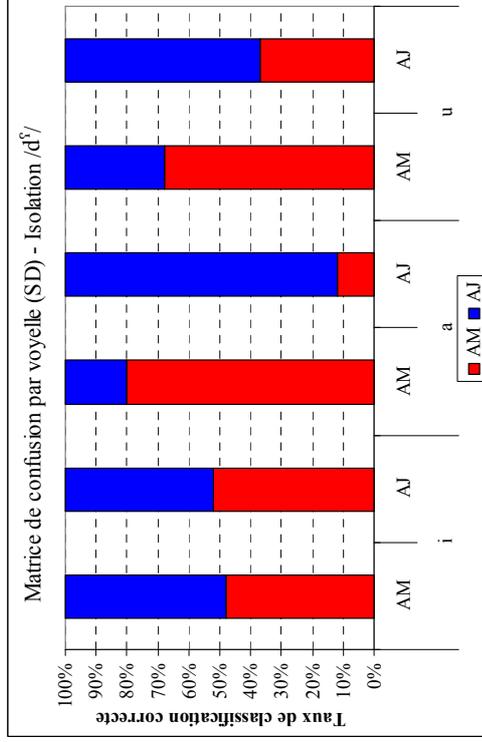
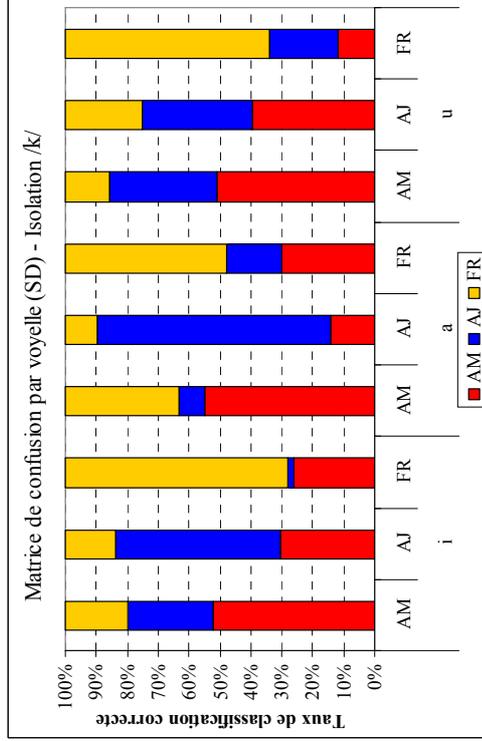
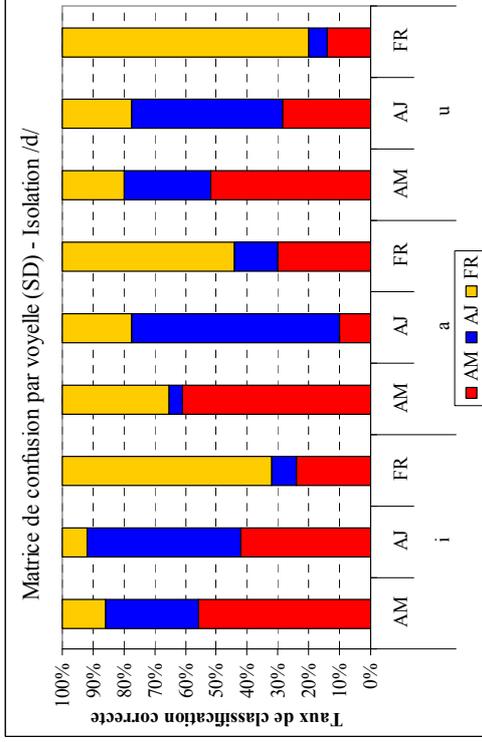
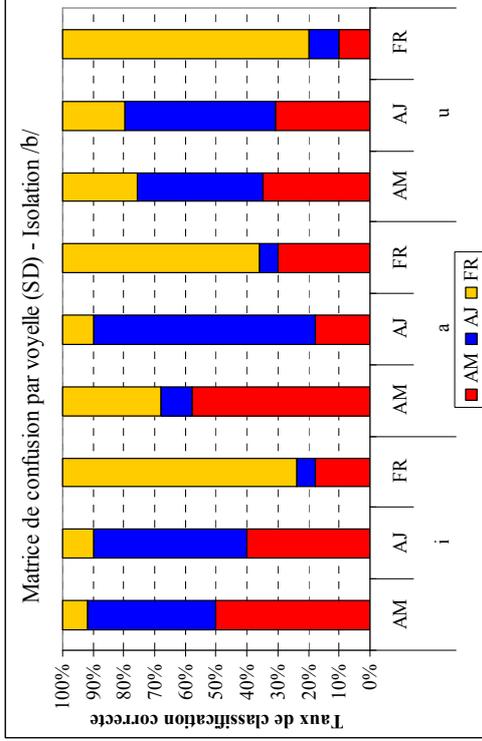
ANNEXE 6 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN STATIQUE AVEC LA DURÉE VOCALIQUE (AD) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.



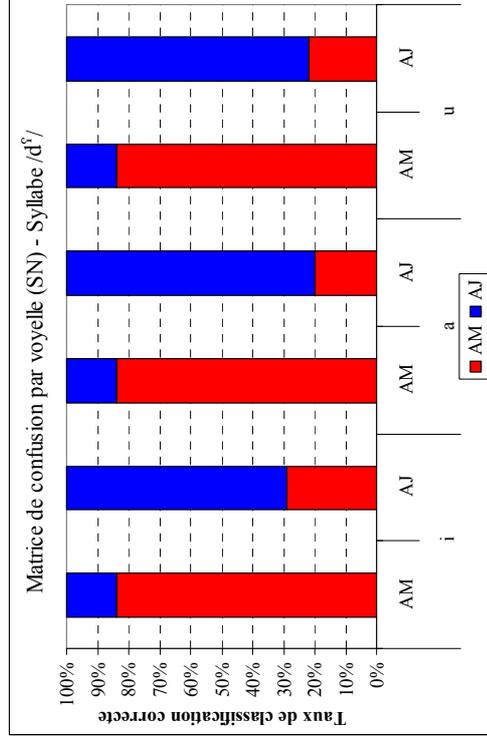
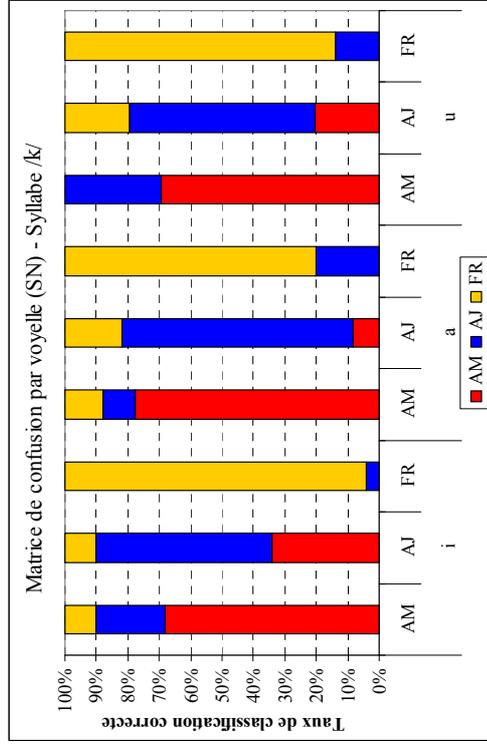
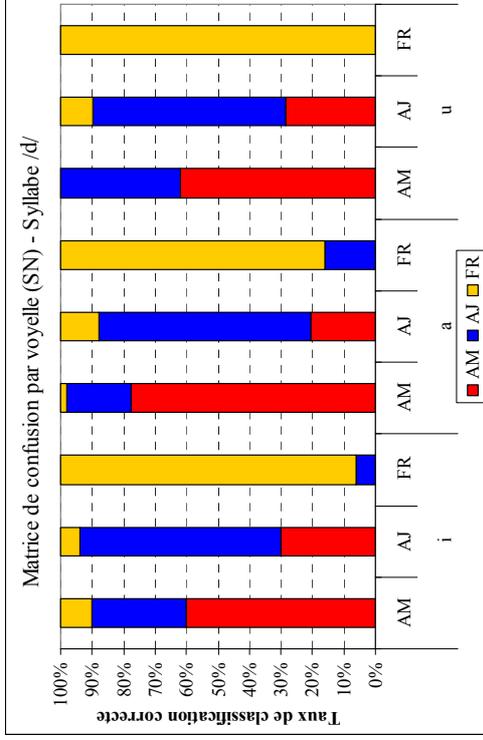
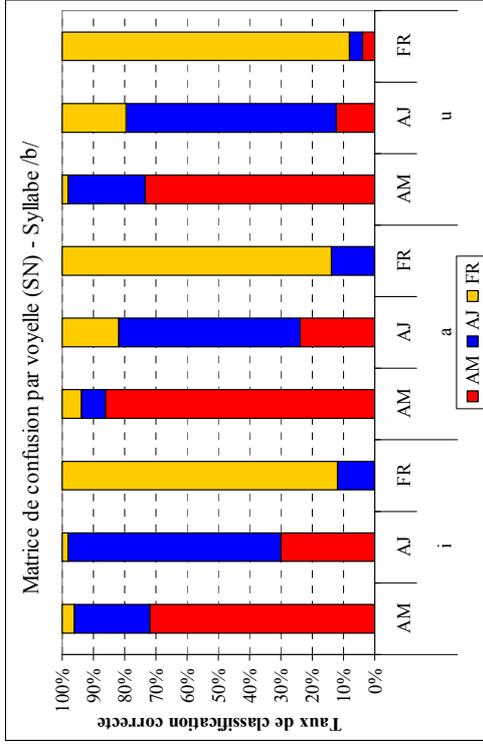


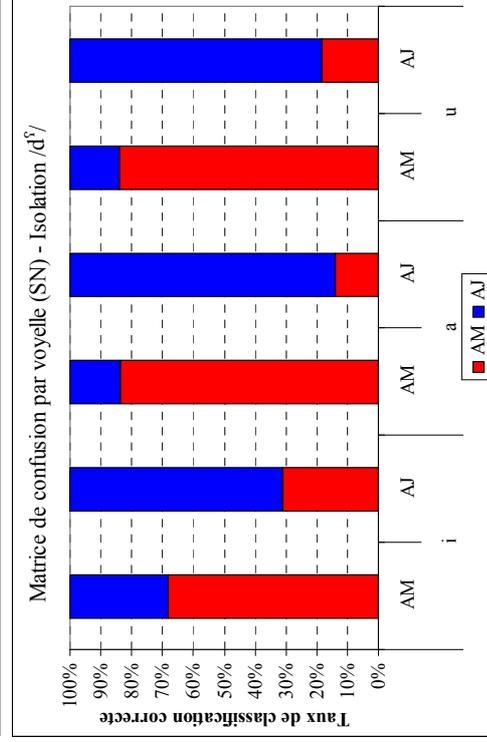
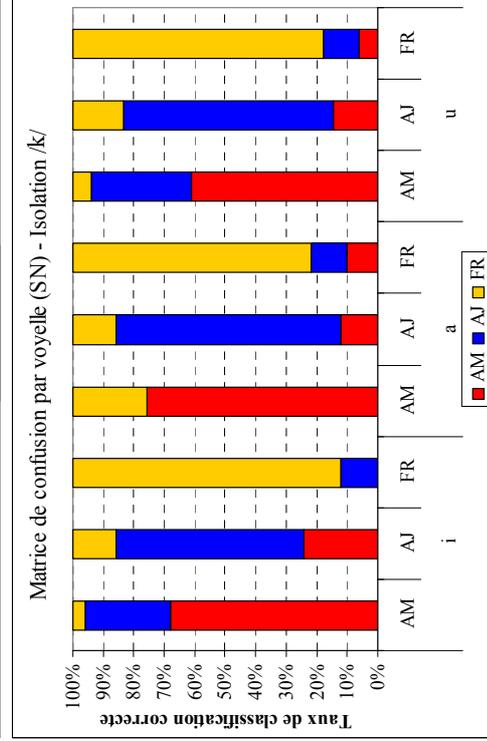
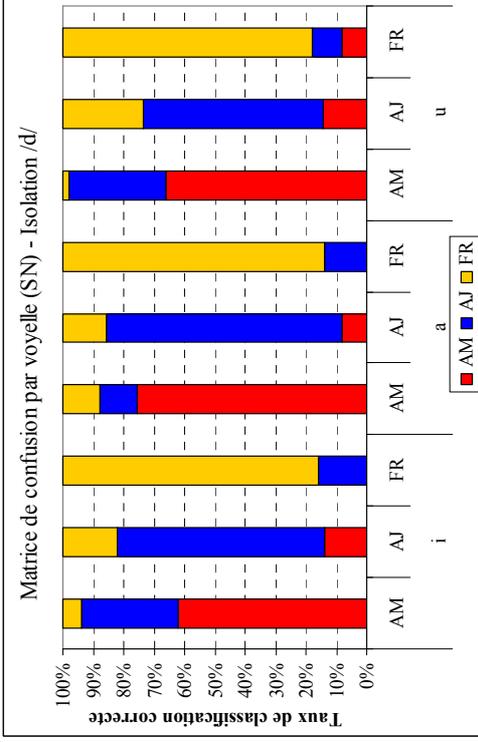
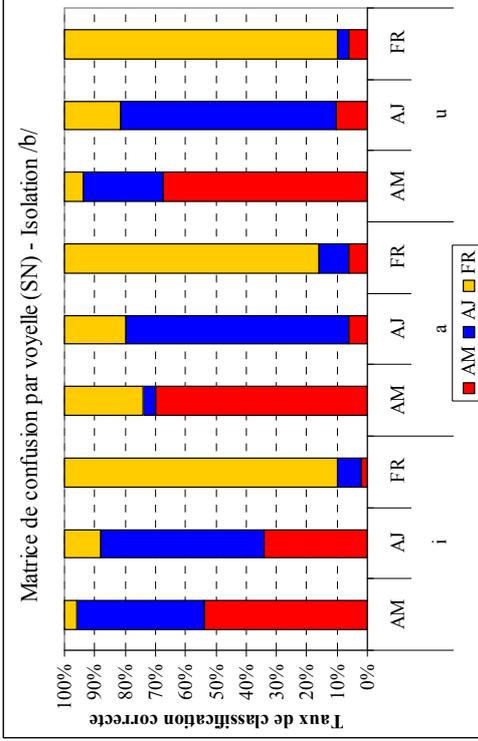
ANNEXE 7 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN STATIQUE SANS LA DURÉE VOCALIQUE (SD) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.



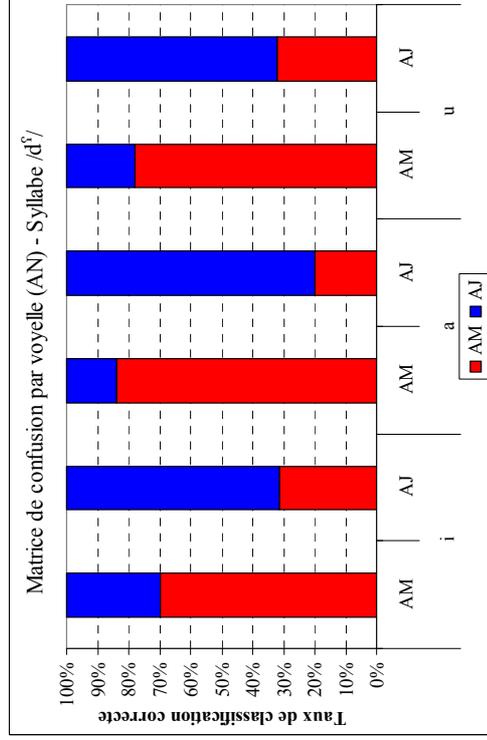
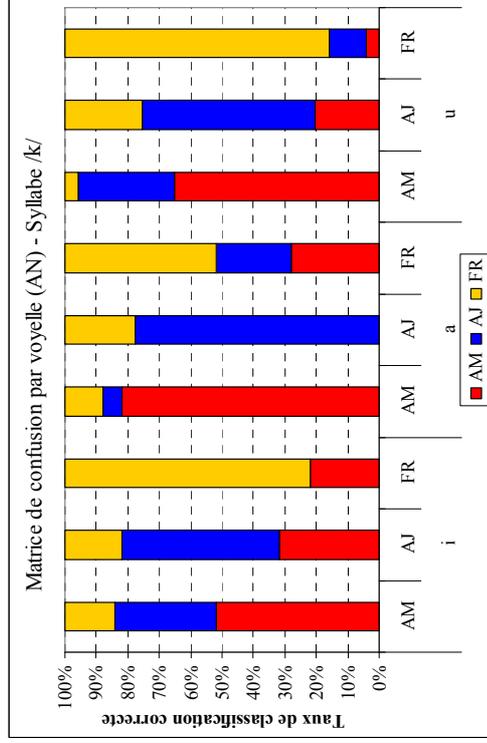
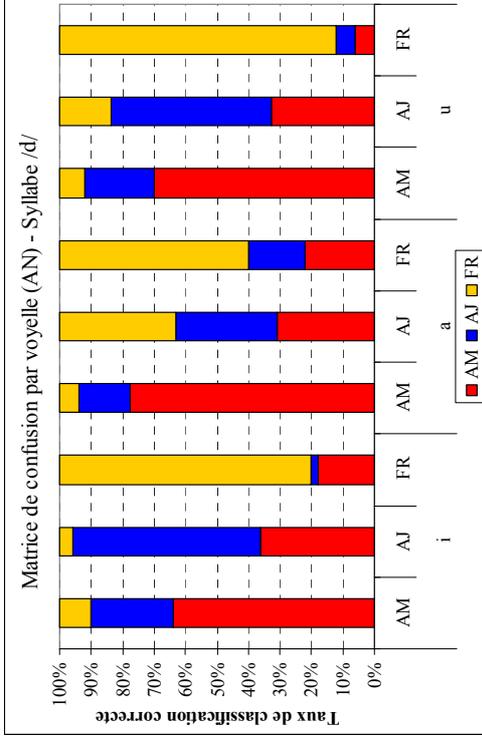
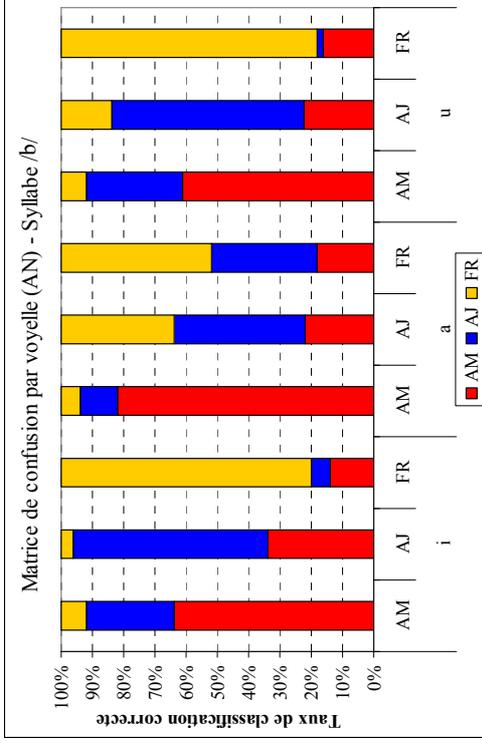


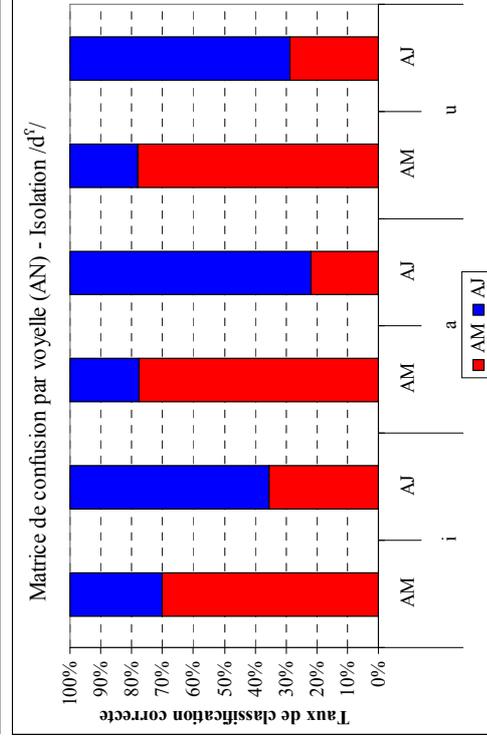
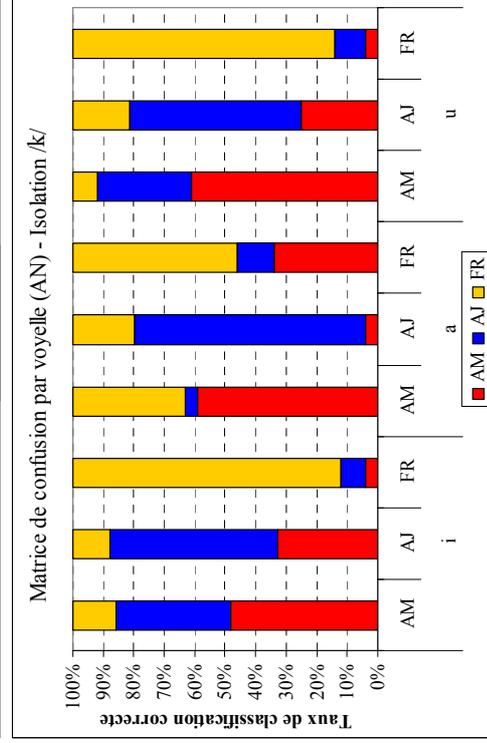
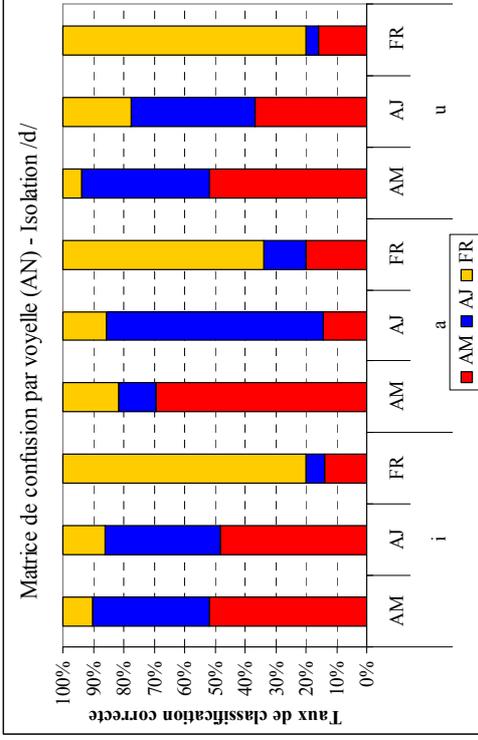
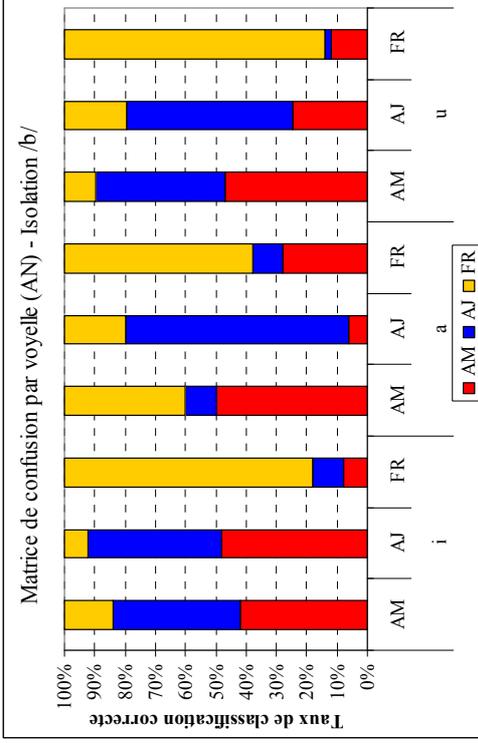
ANNEXE 8 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION LINÉAIRE) SANS NORMALISATION DU TEMPS (SN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.



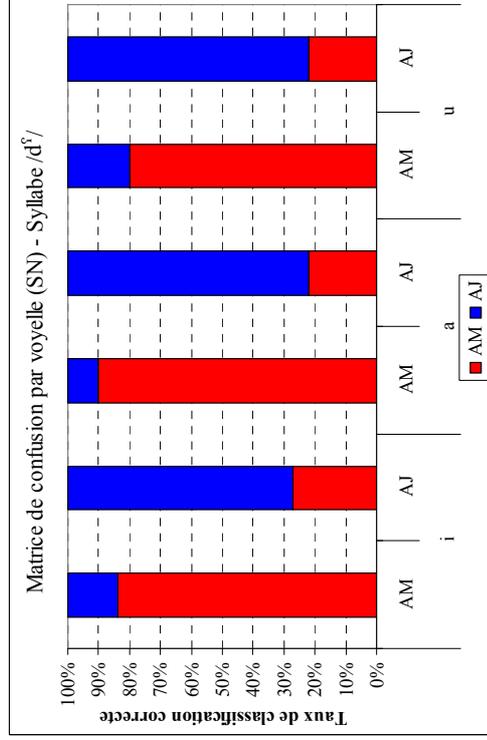
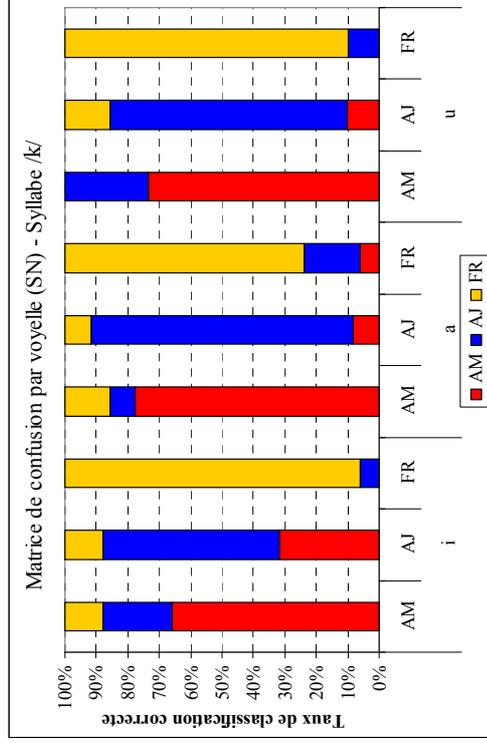
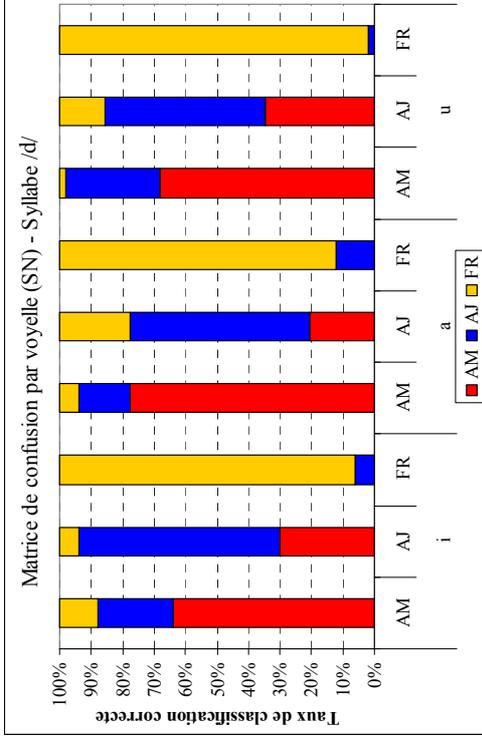
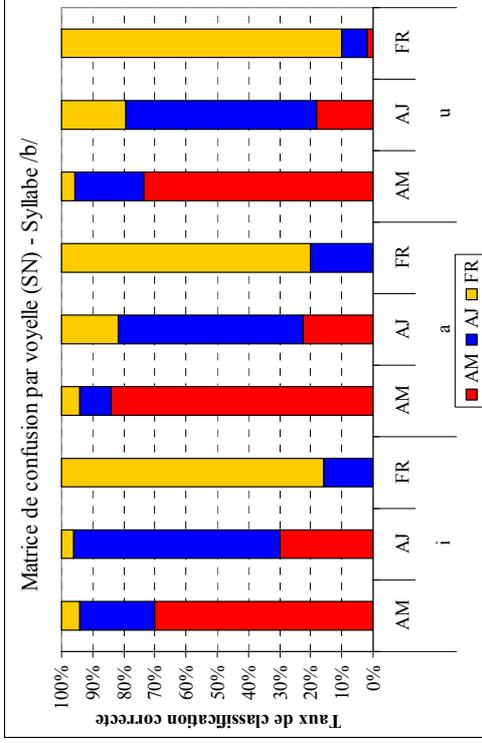


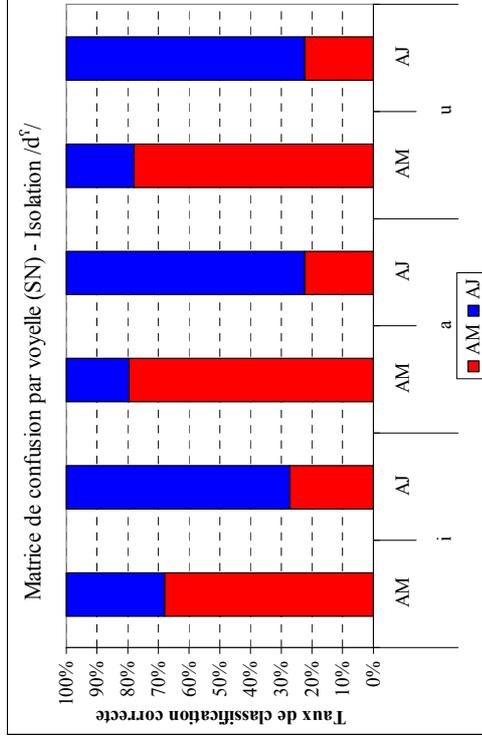
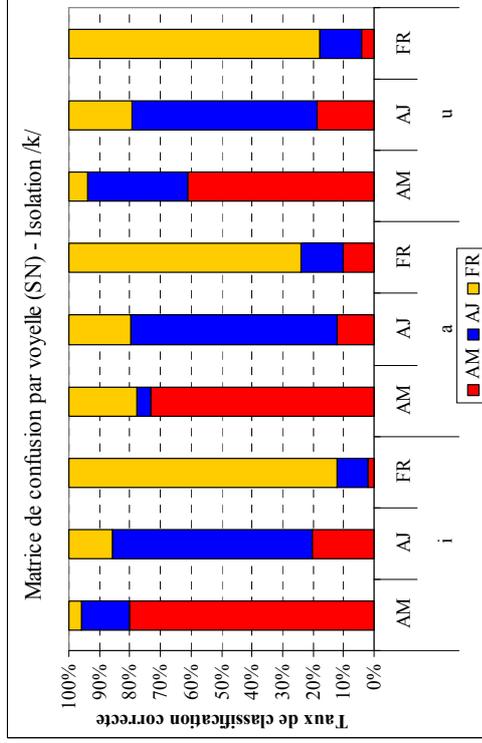
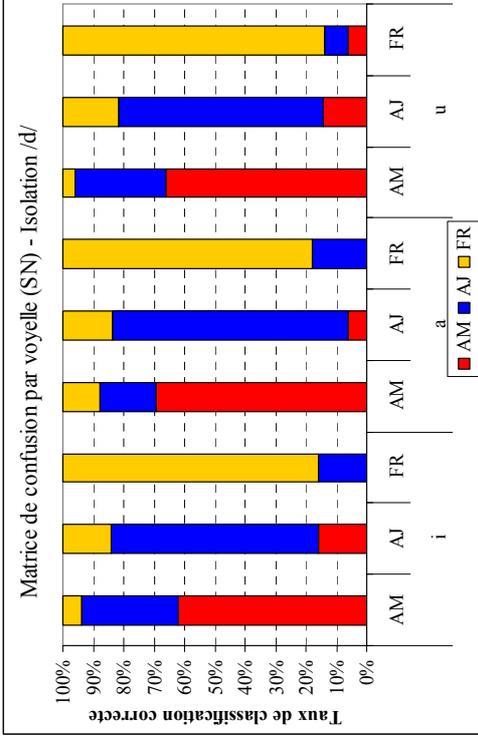
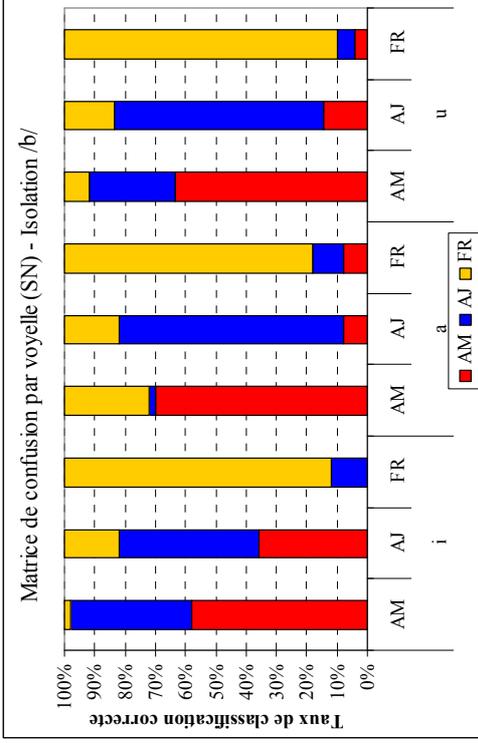
ANNEXE 9 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC NORMALISATION DU TEMPS (AN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.



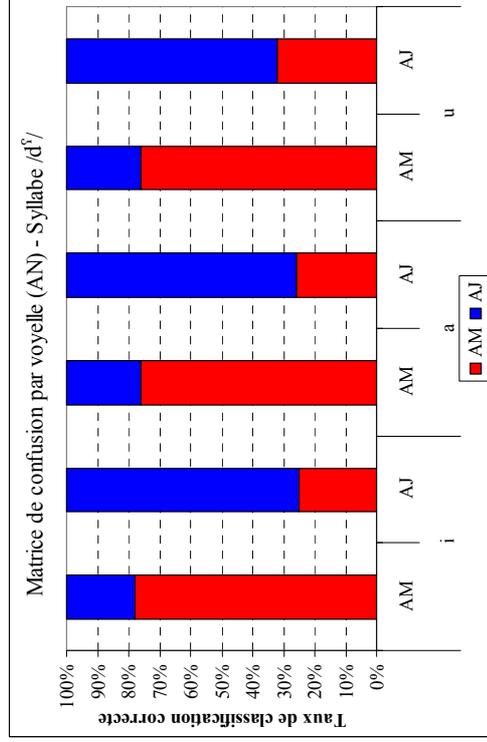
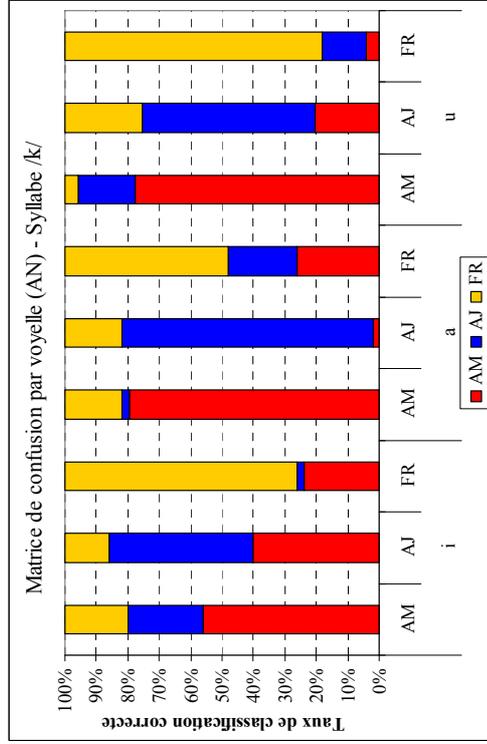
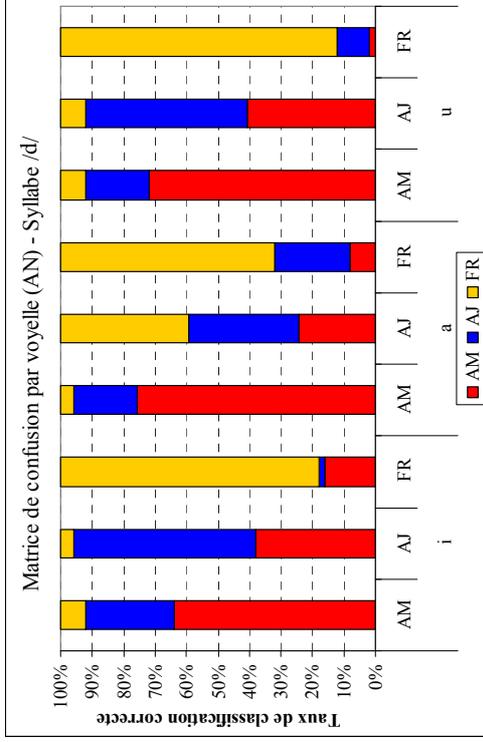
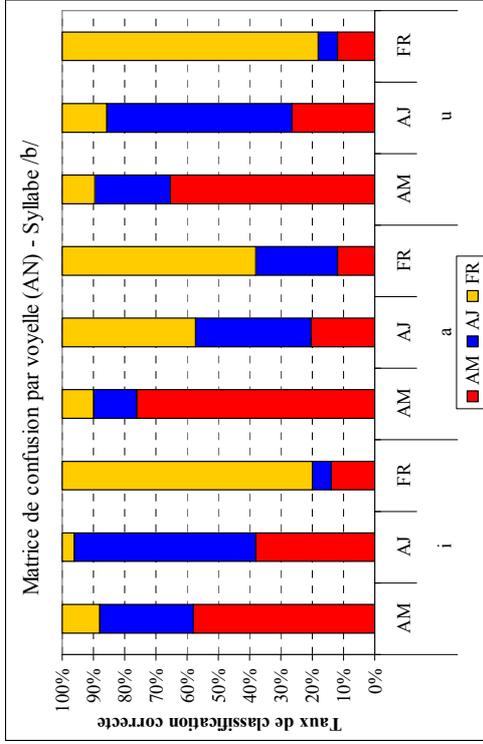


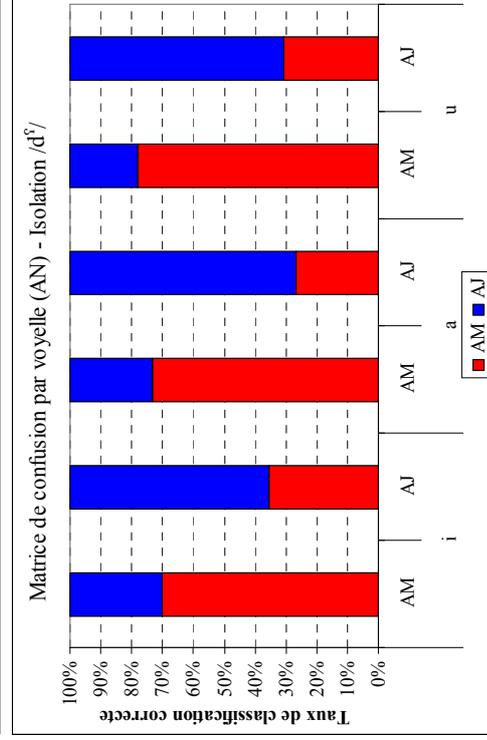
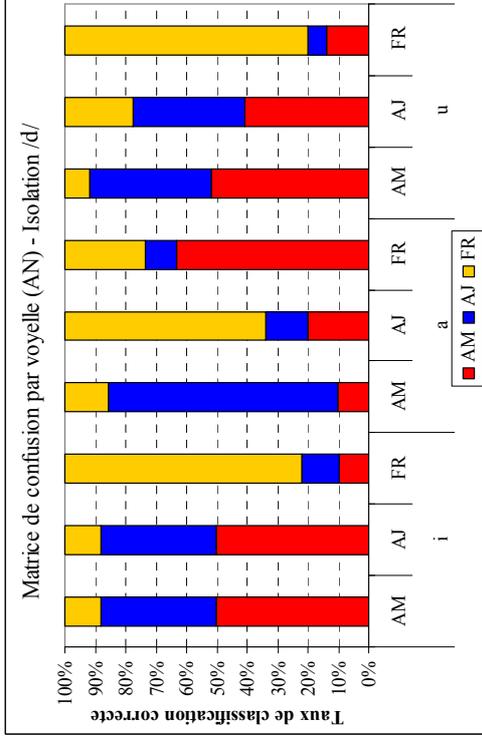
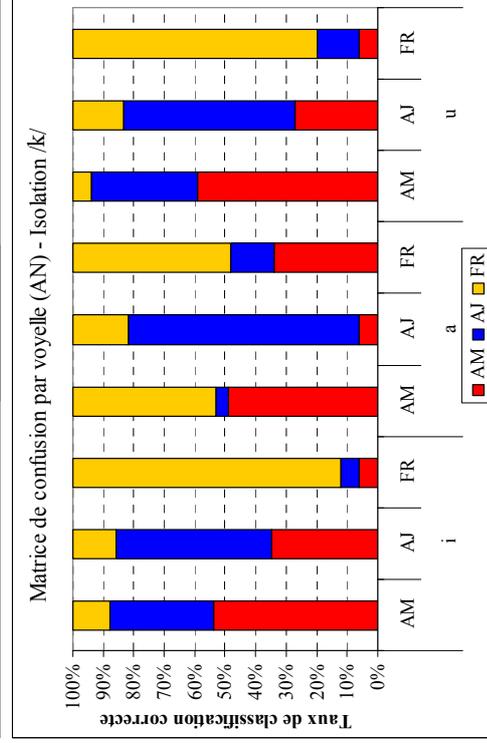
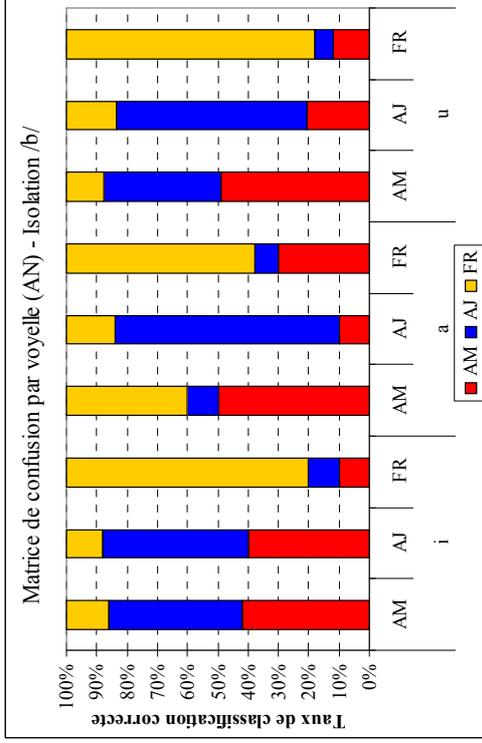
ANNEXE 10 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2^{ÈME} ORDRE) SANS NORMALISATION DU TEMPS (SN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.



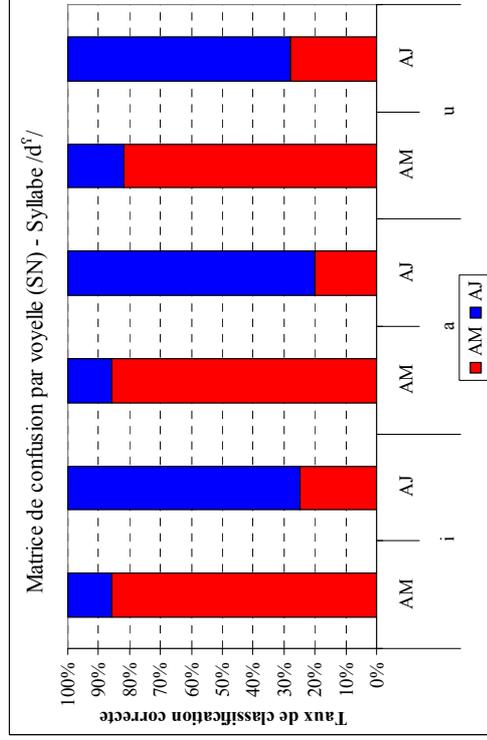
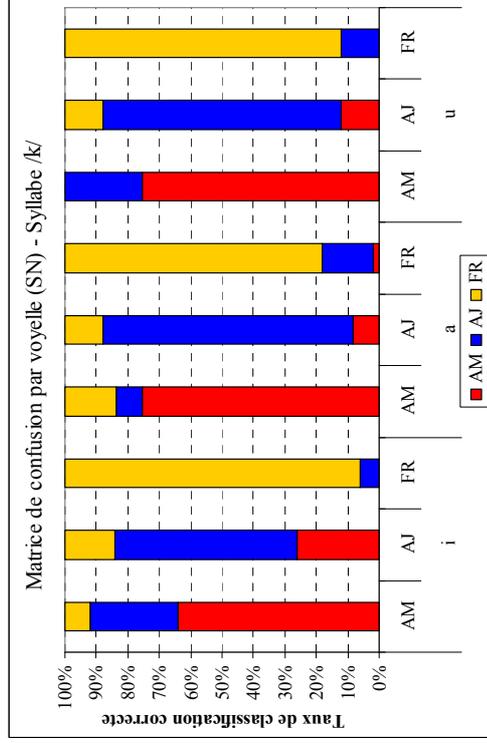
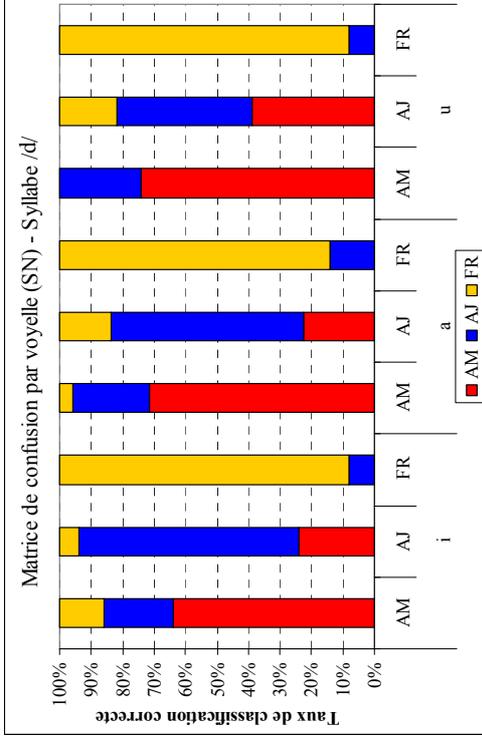
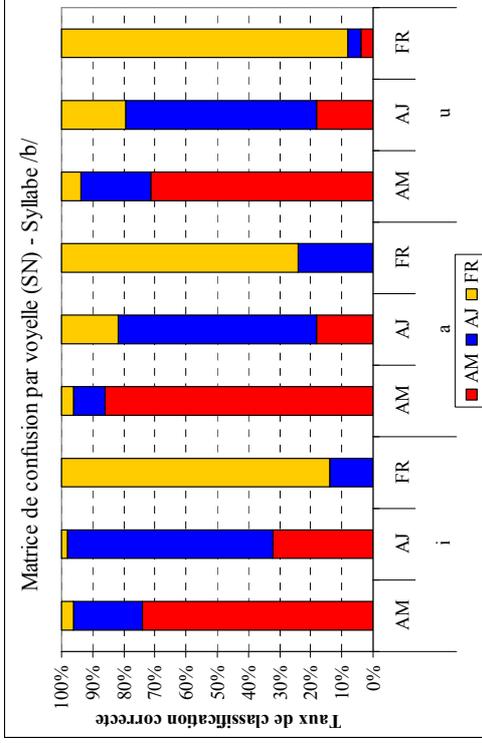


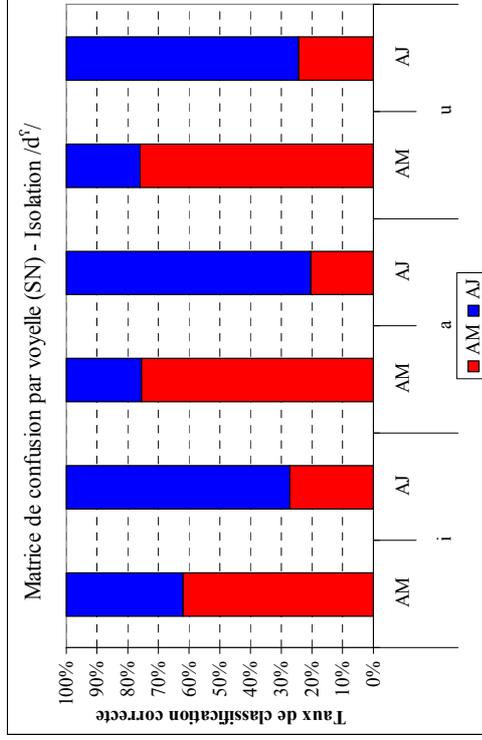
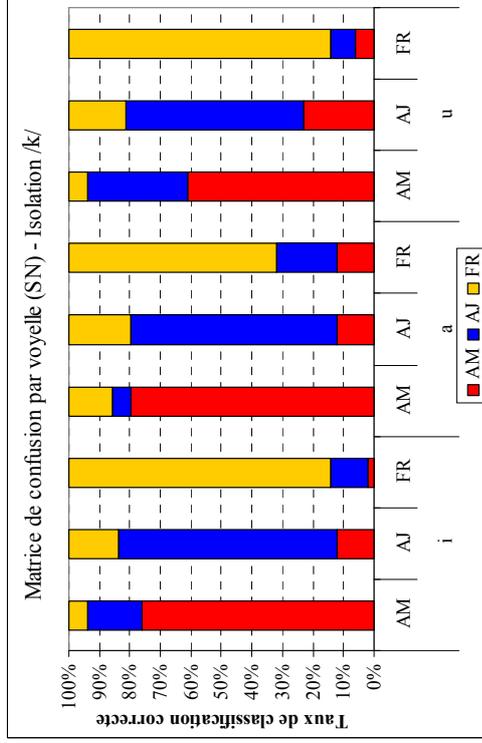
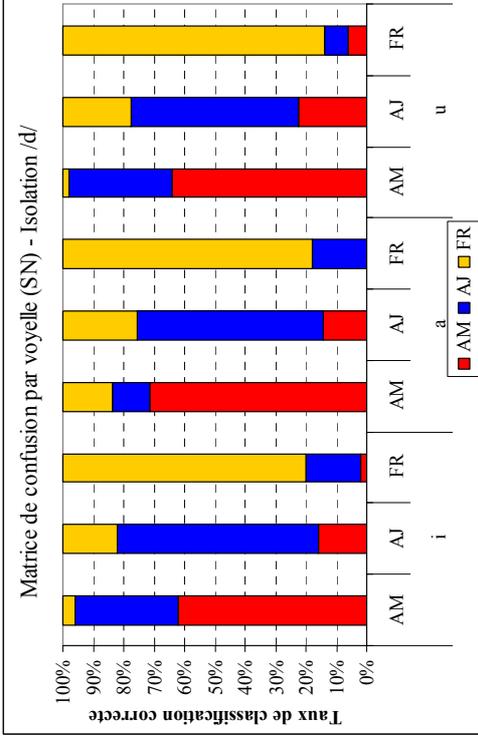
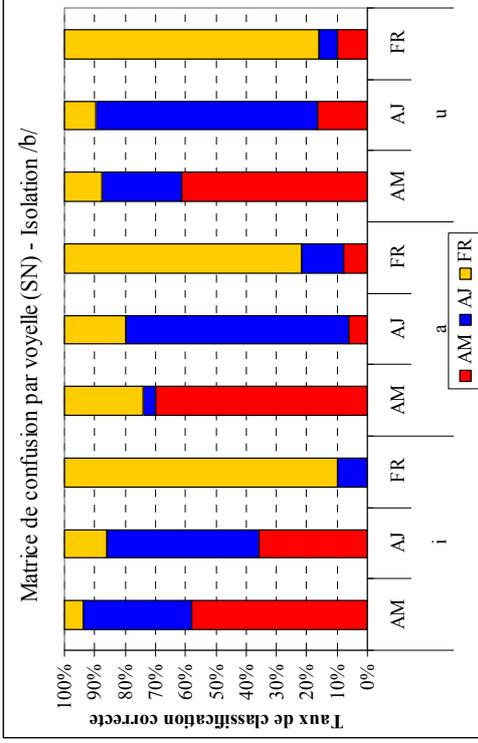
ANNEXE 11 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2^{ÈME} ORDRE) AVEC NORMALISATION DU TEMPS (AN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.





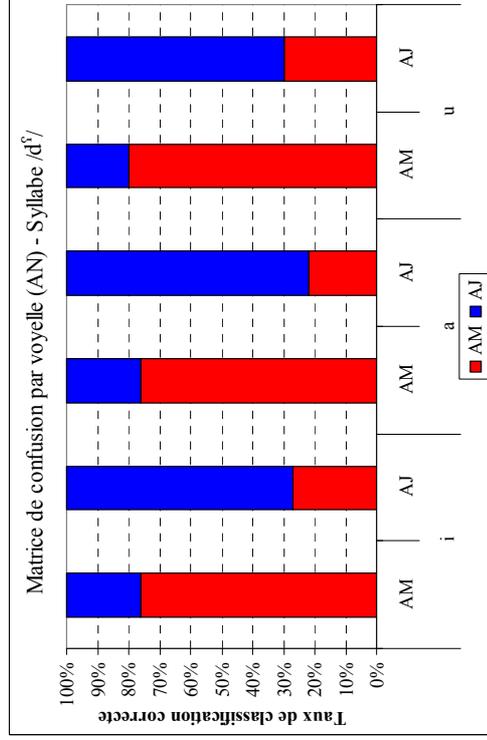
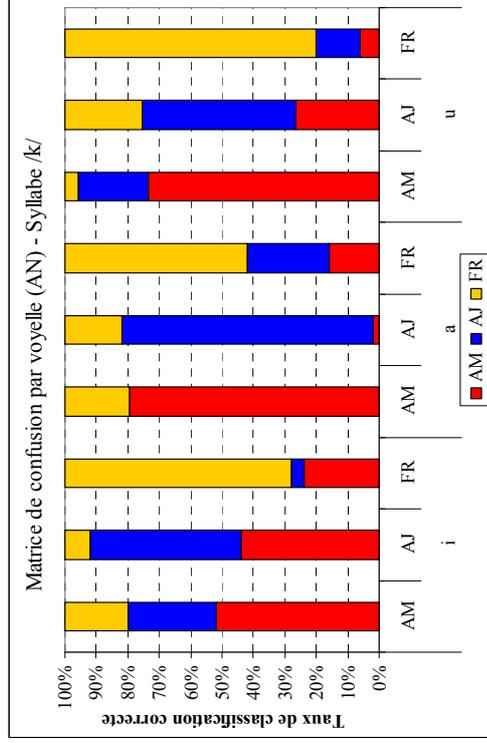
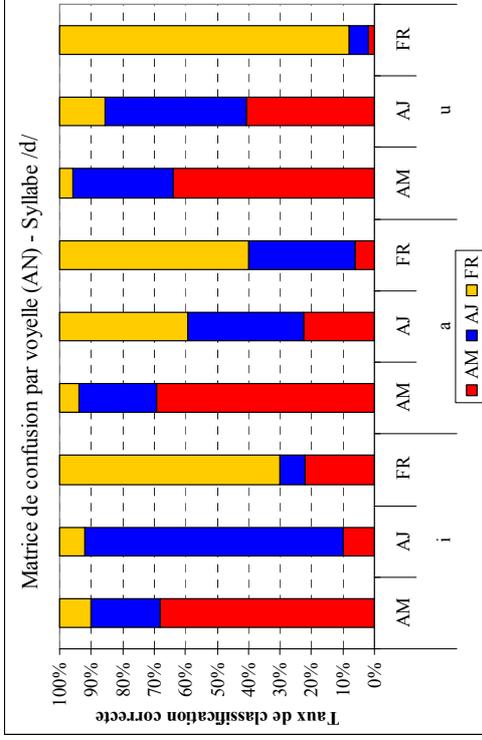
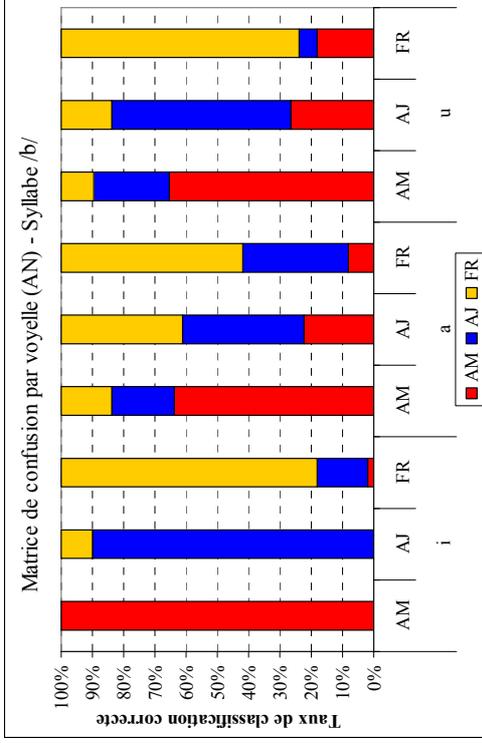
ANNEXE 12 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3^{ÈME} ORDRE) SANS NORMALISATION DU TEMPS (SN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.

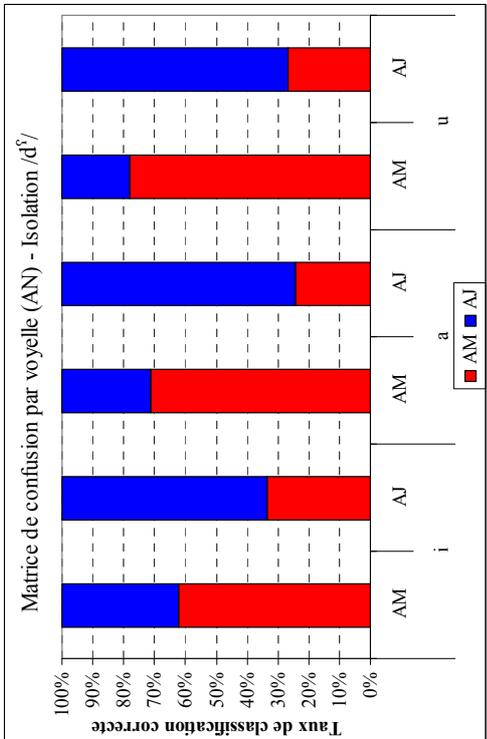
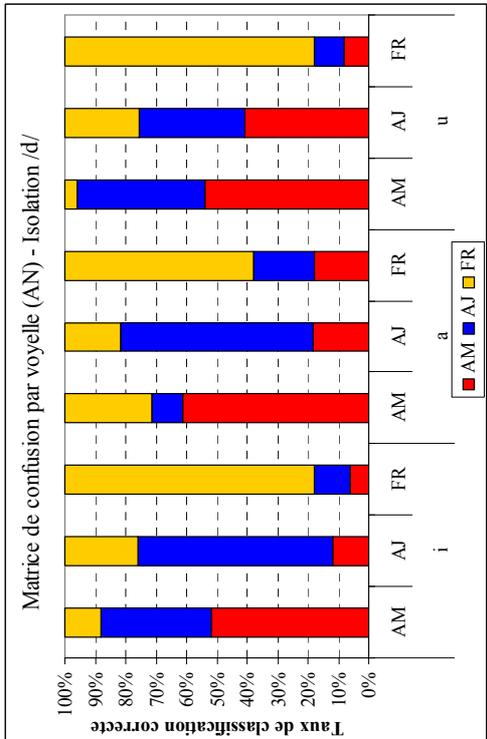
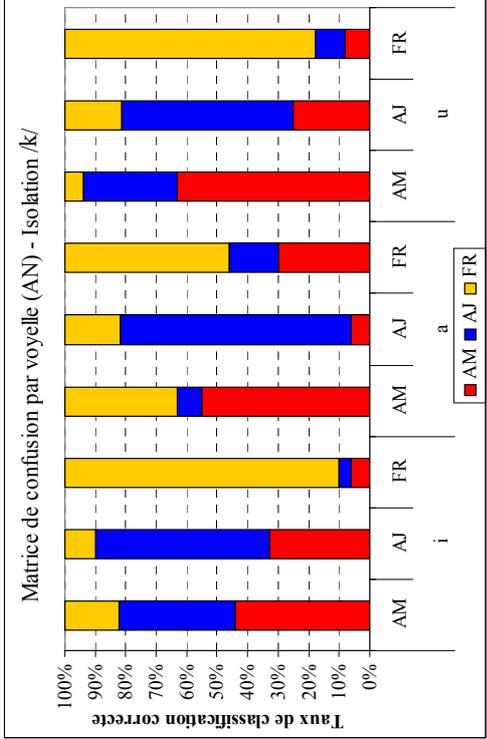
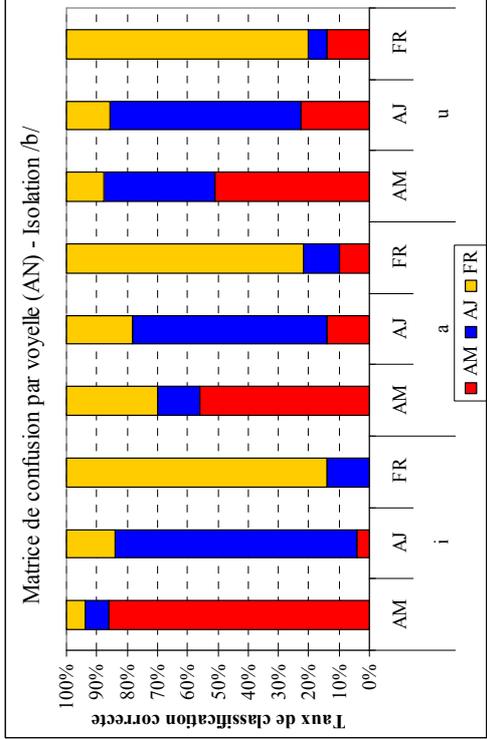




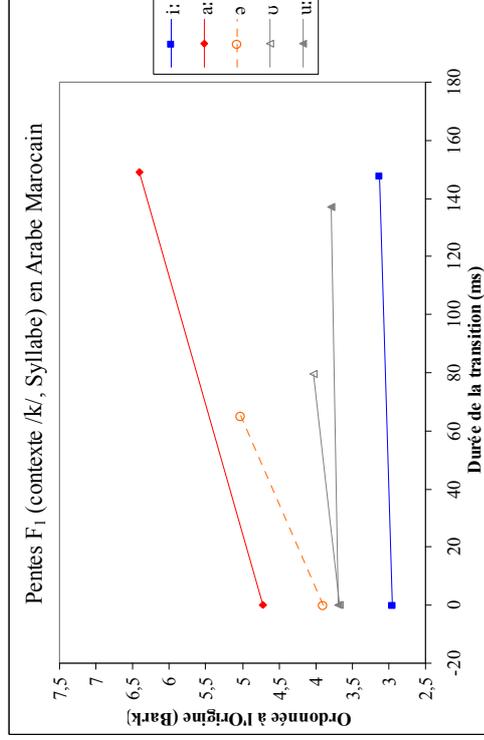
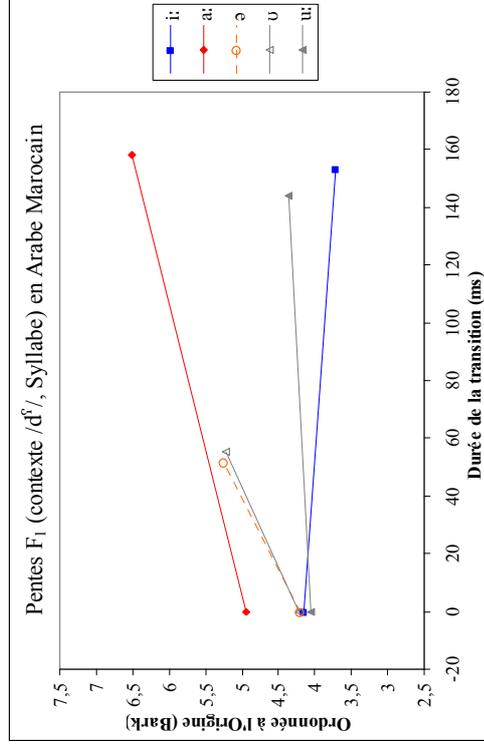
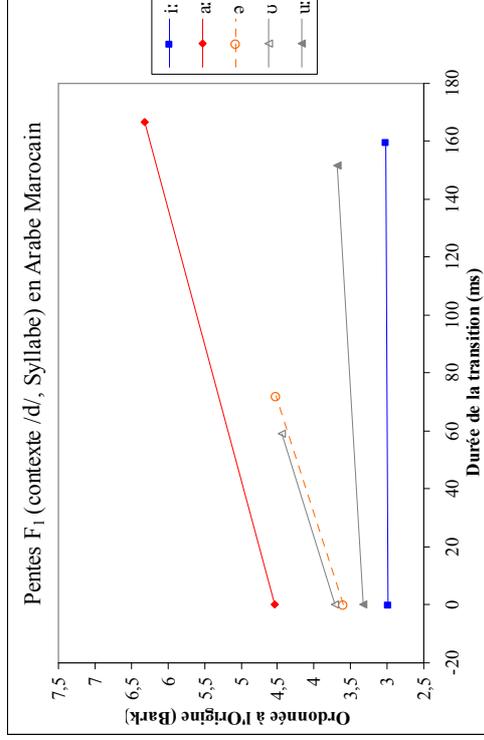
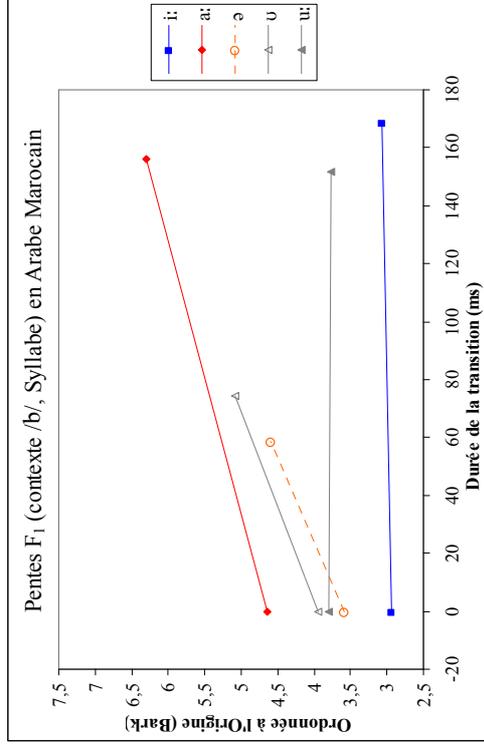
ANNEXE 13

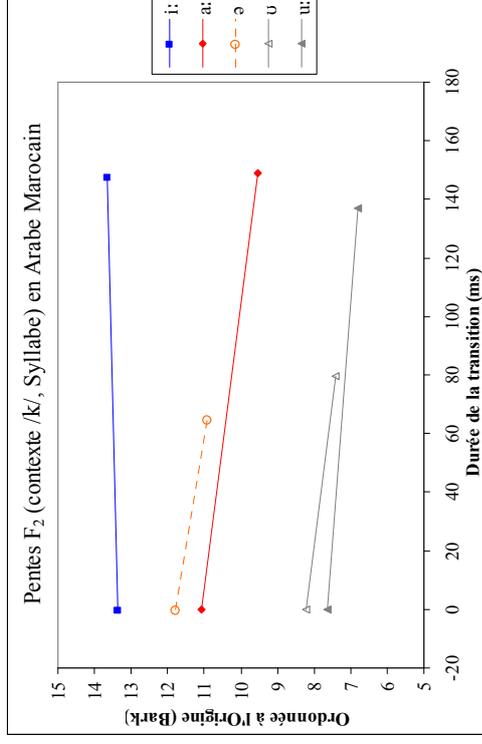
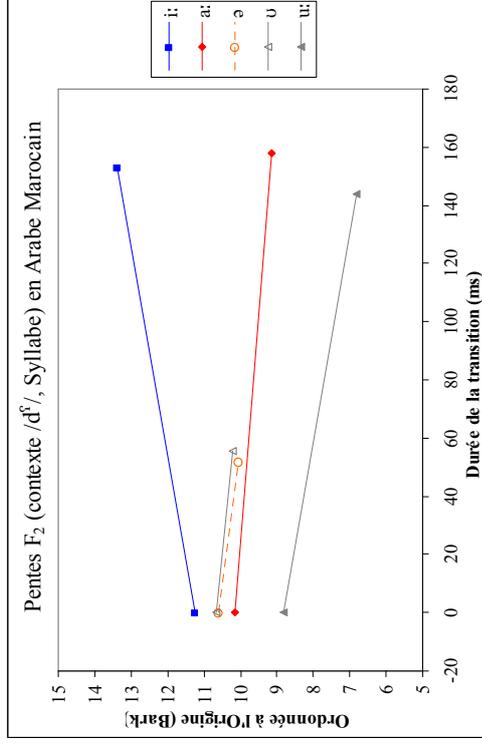
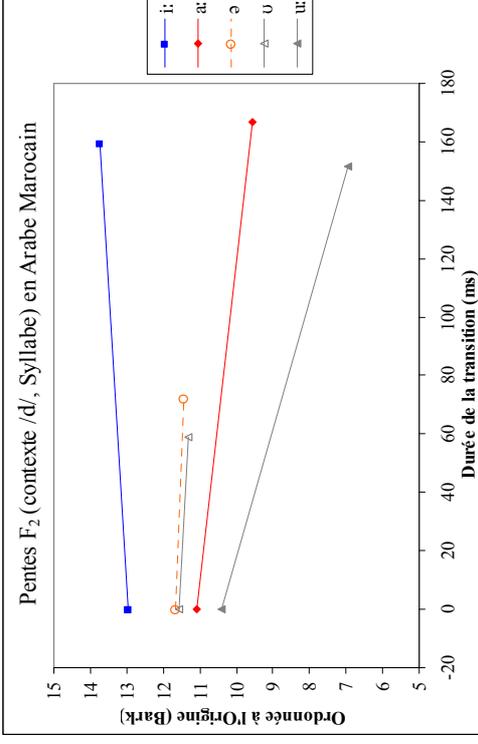
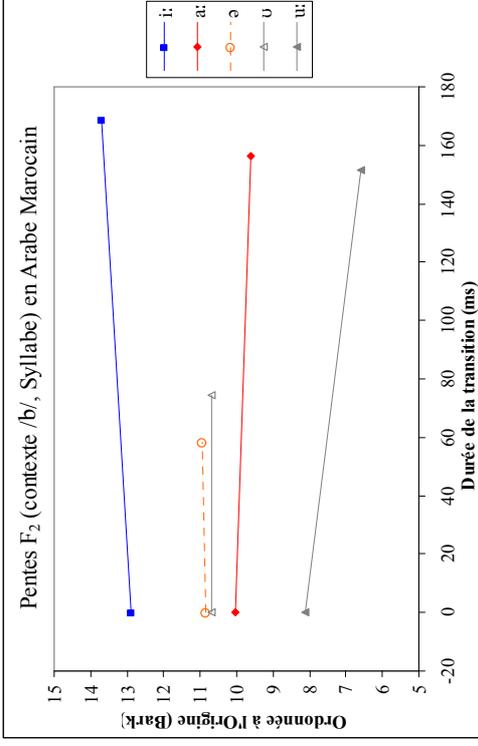
MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES /i a u/ EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3^{ÈME} ORDRE) AVEC NORMALISATION DU TEMPS (AN) DANS LES TROIS LANGUES EN SITUATIONS SYLLABE ET ISOLATION.

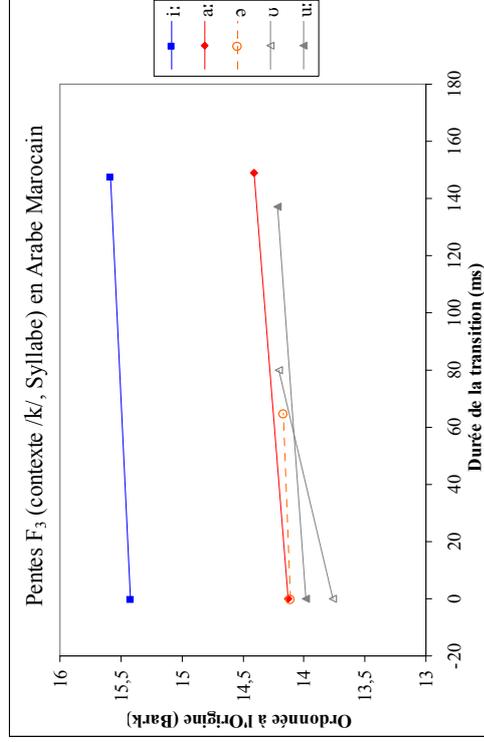
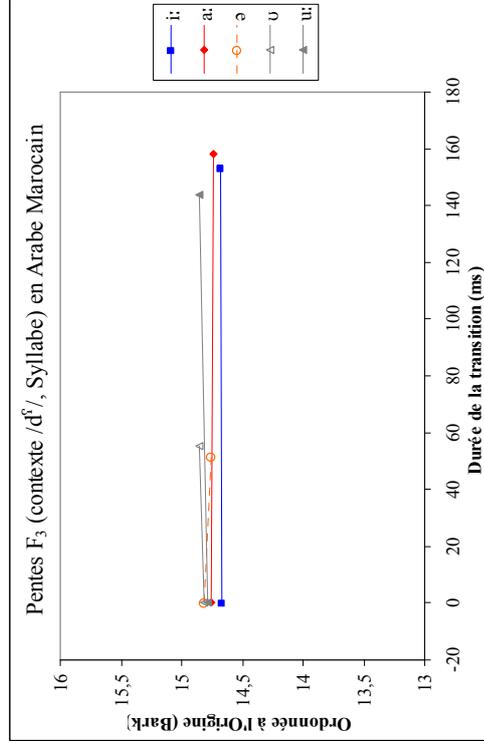
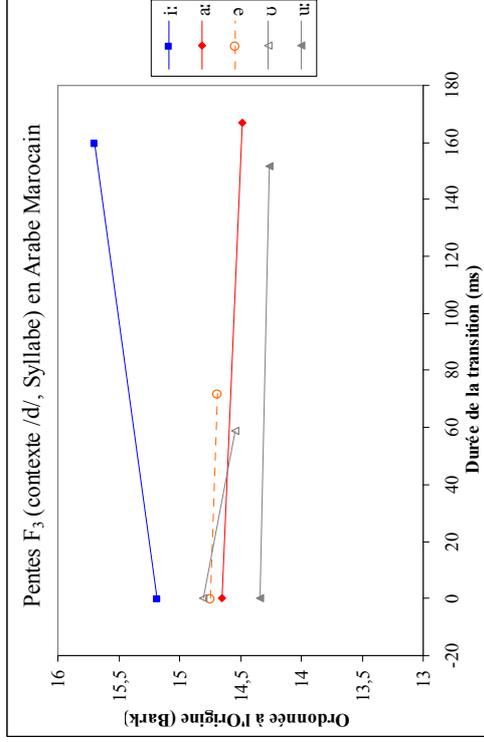
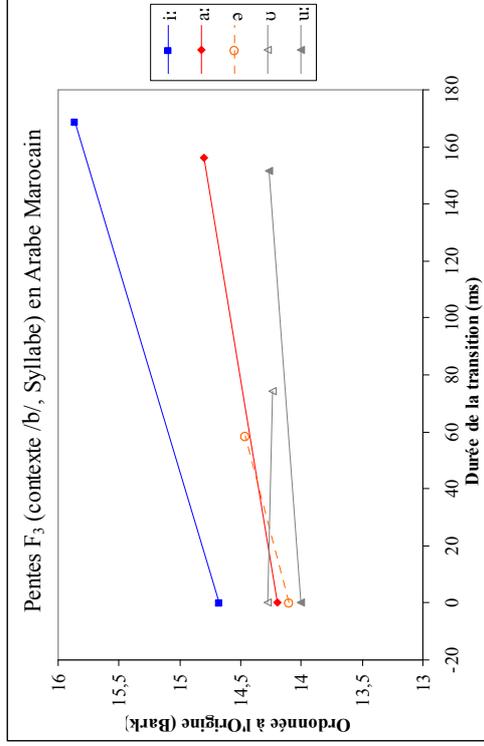


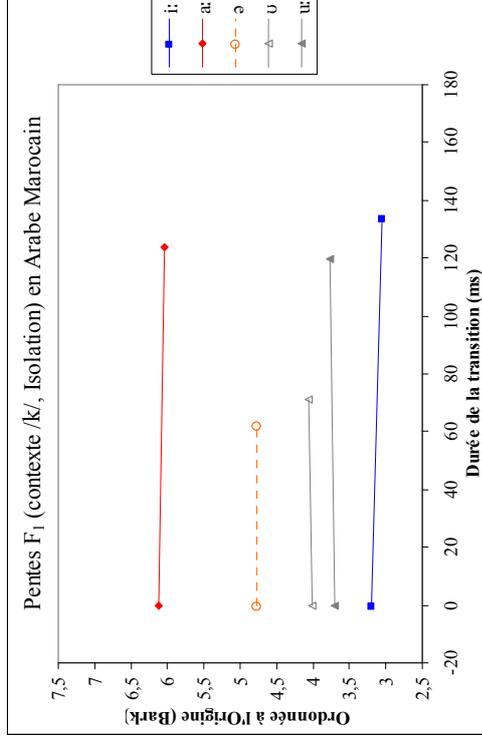
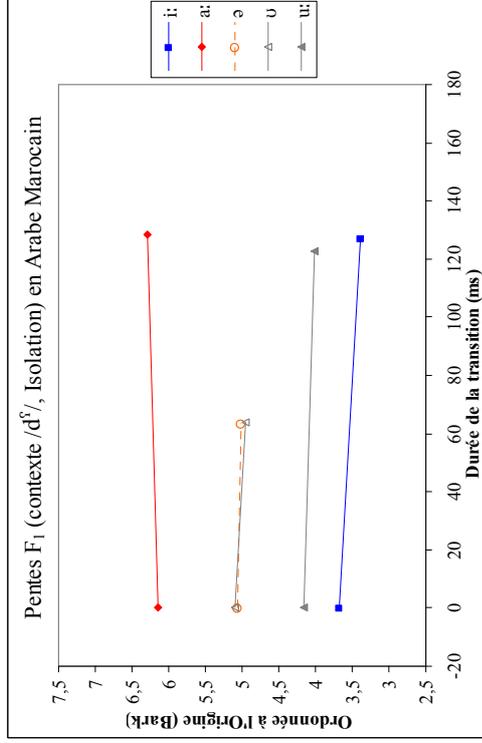
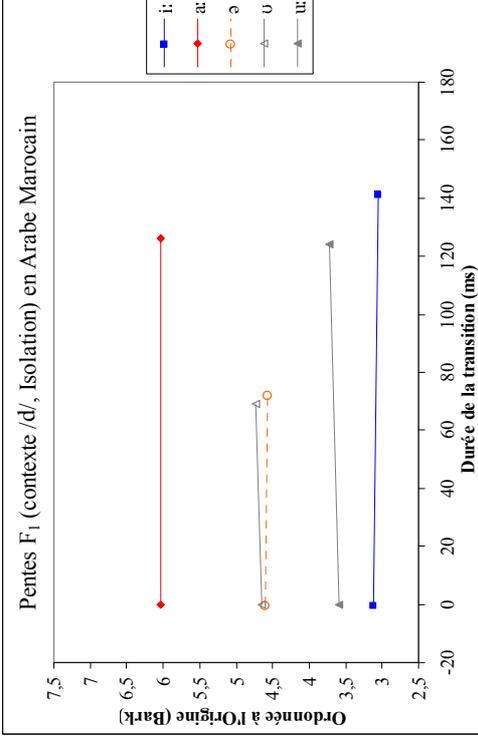
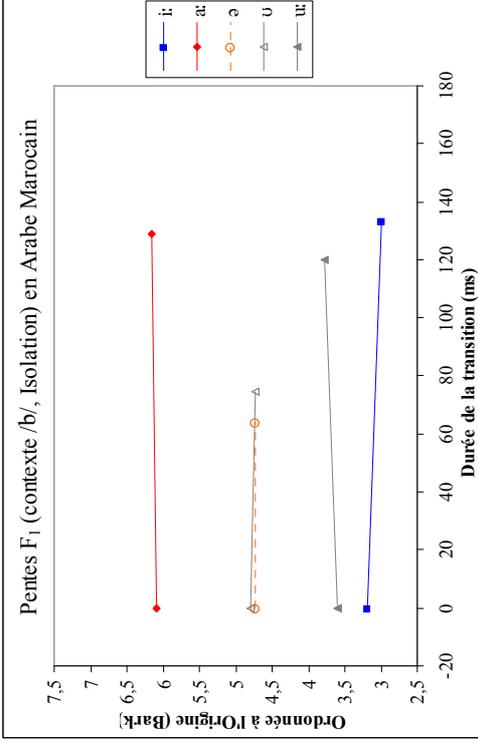


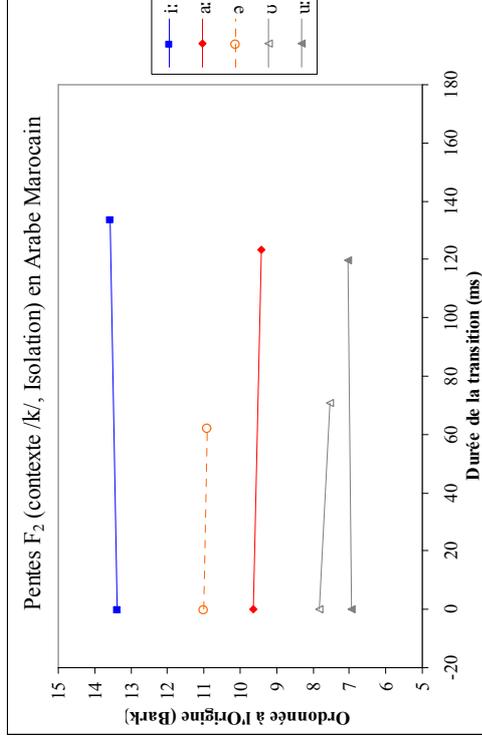
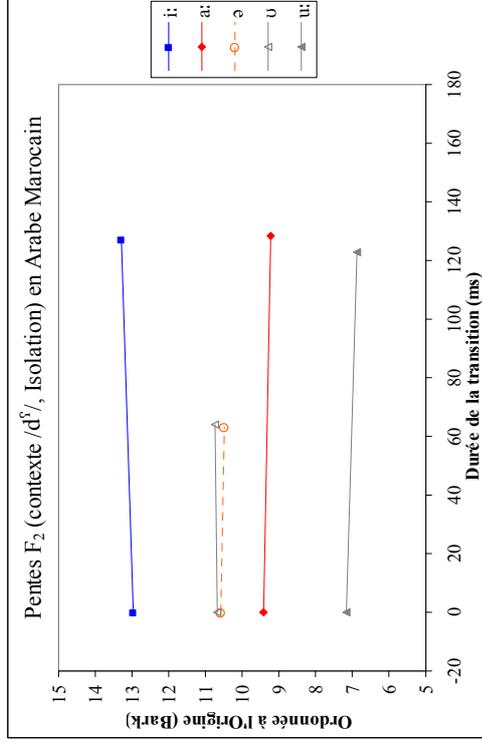
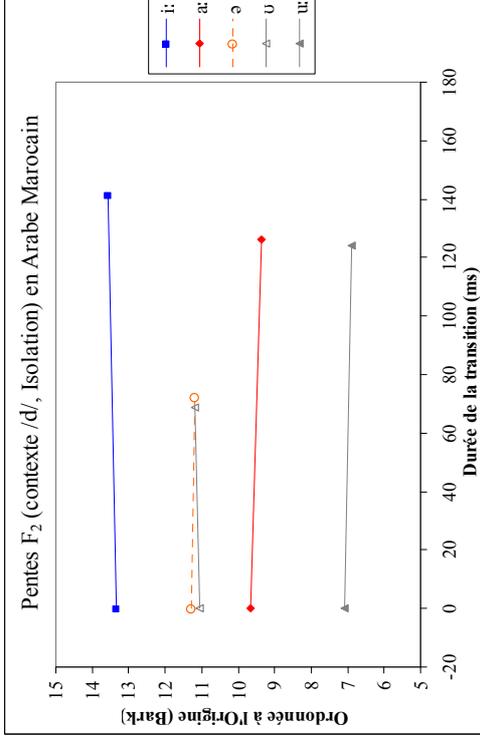
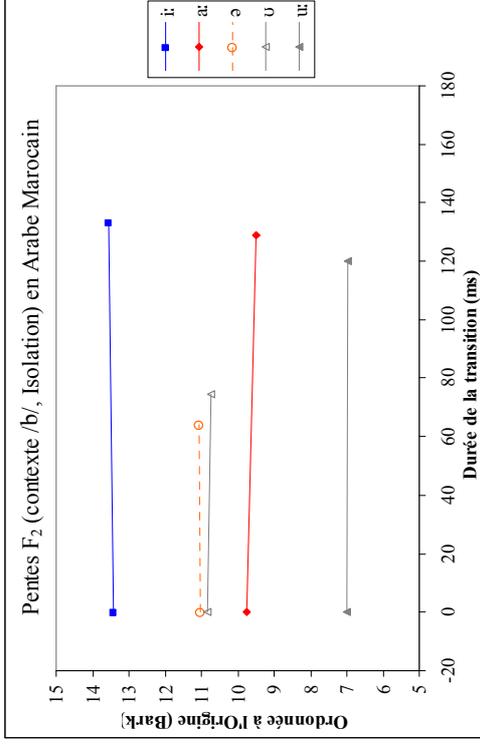
ANNEXE 14 PENTES FORMANTIQUES DE F₁, F₂ & F₃ PAR VOYELLE EN SITUATIONS SYLLABE & ISOLATION EN ARABE MAROCAIN

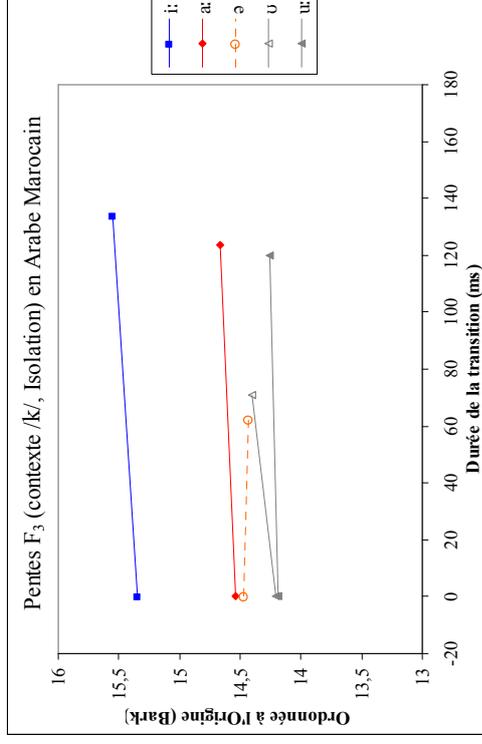
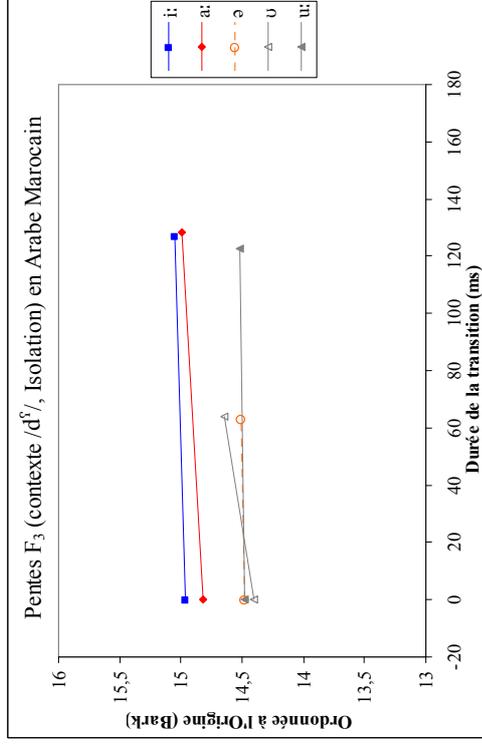
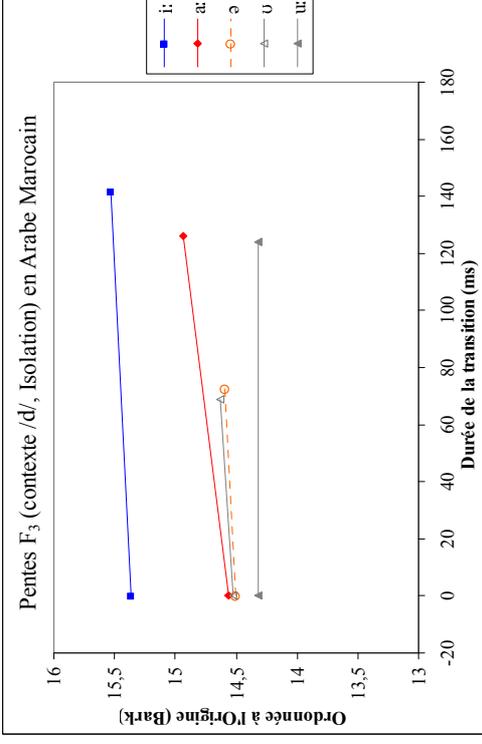
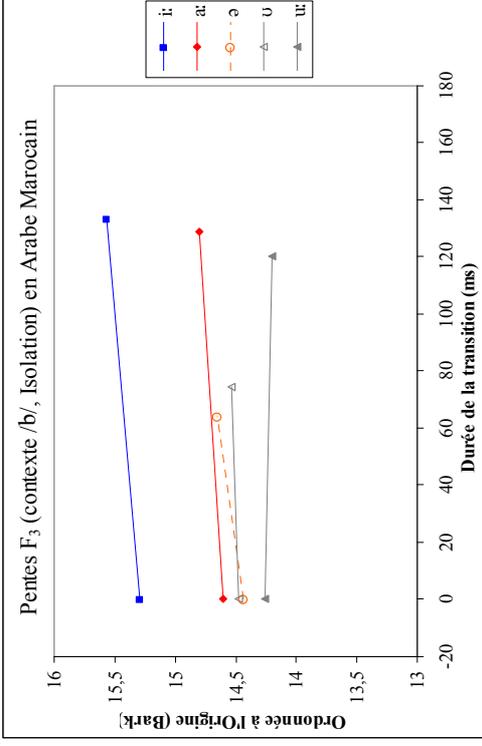




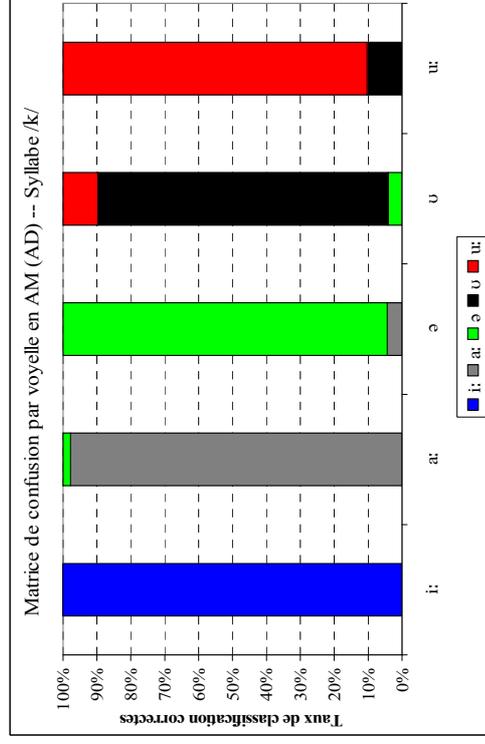
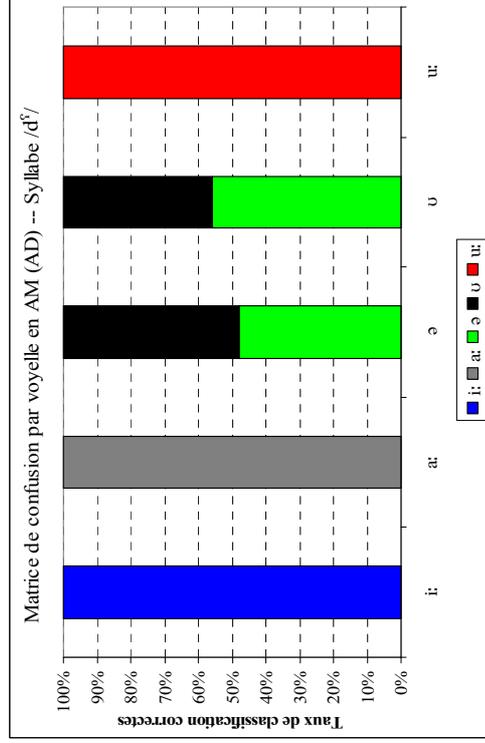
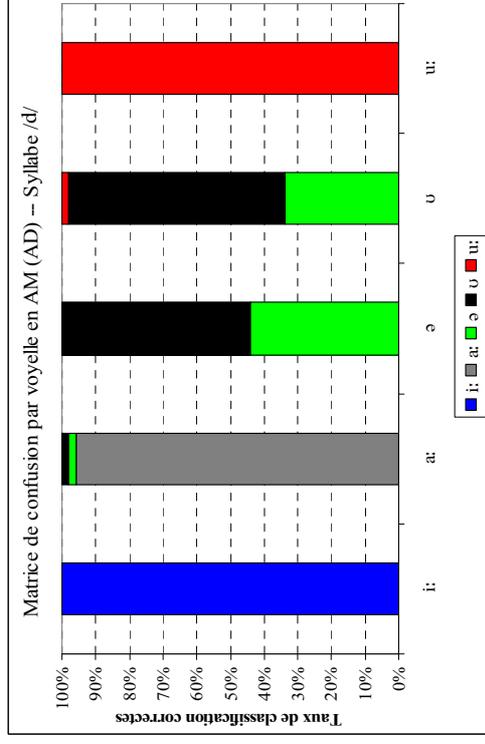
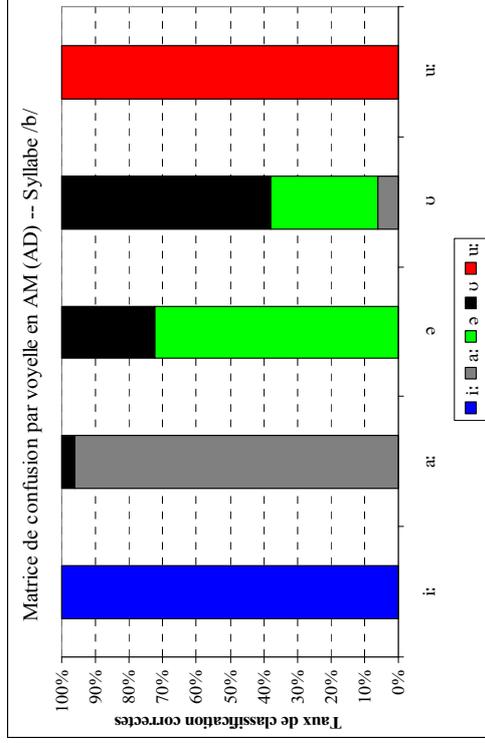


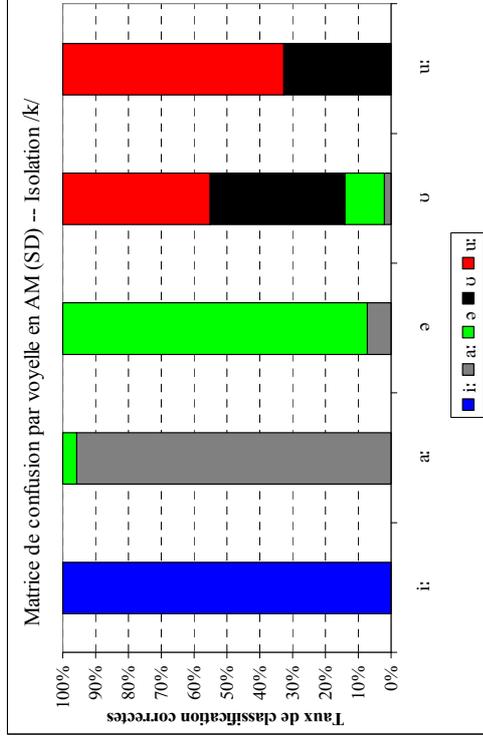
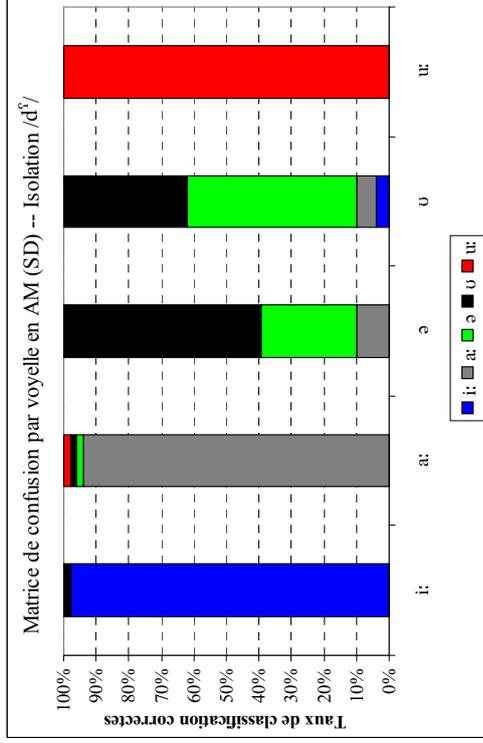
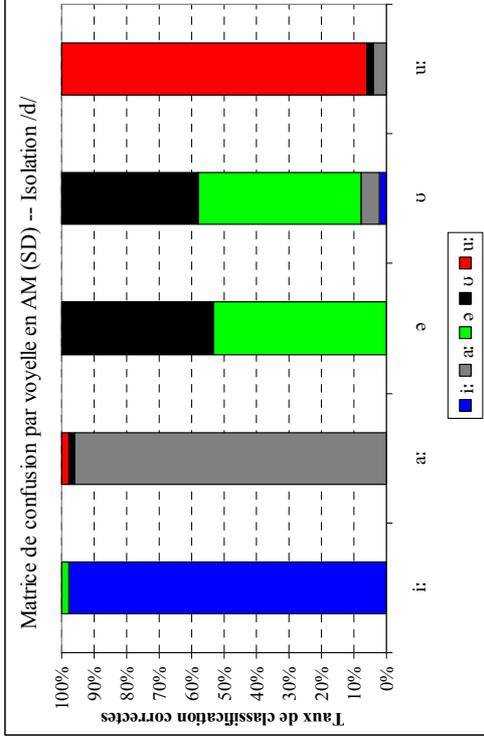
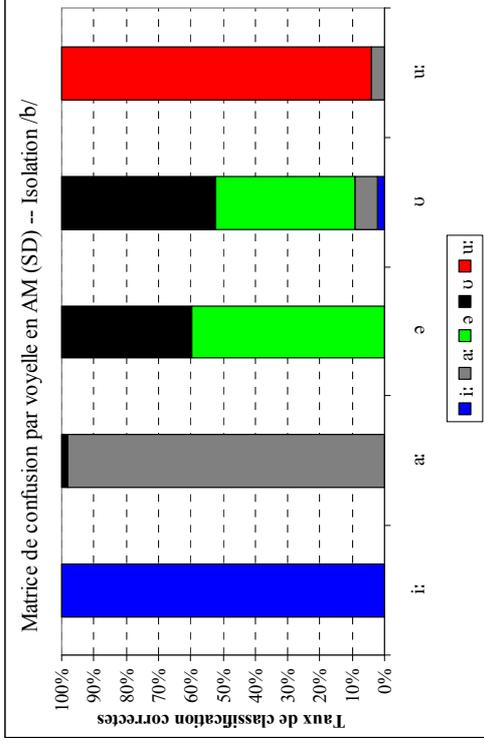




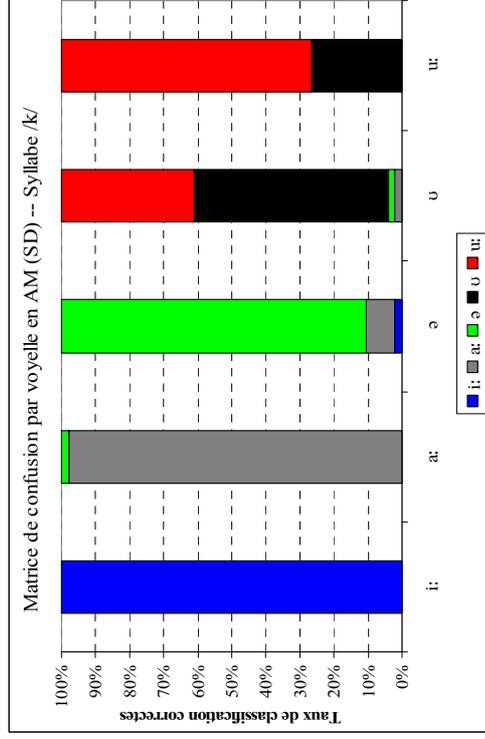
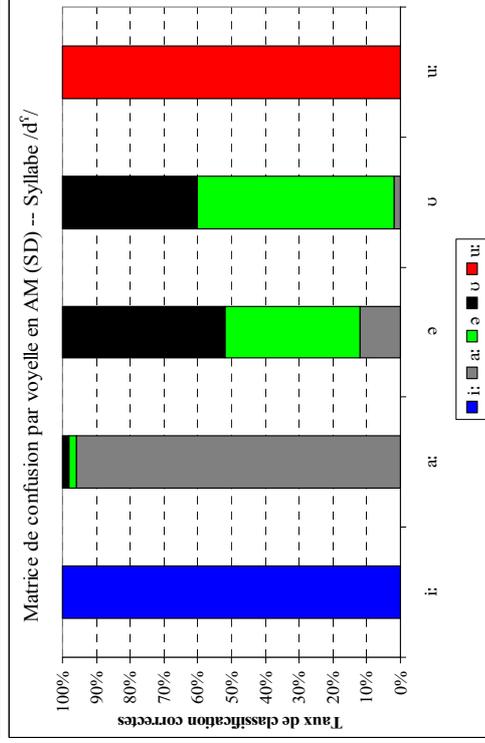
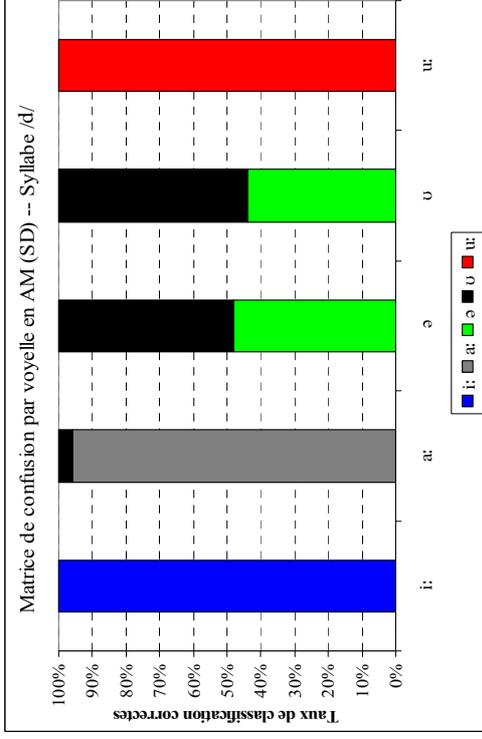
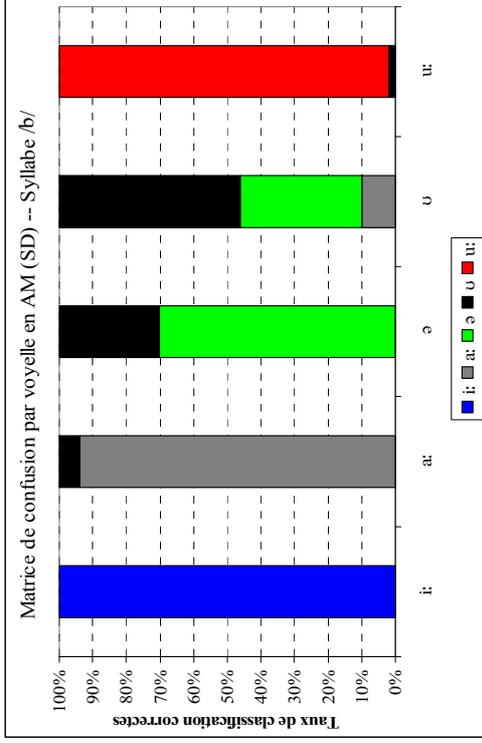


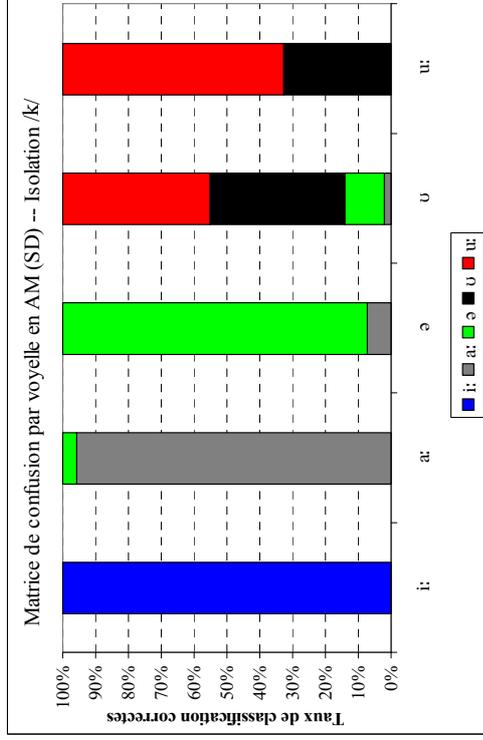
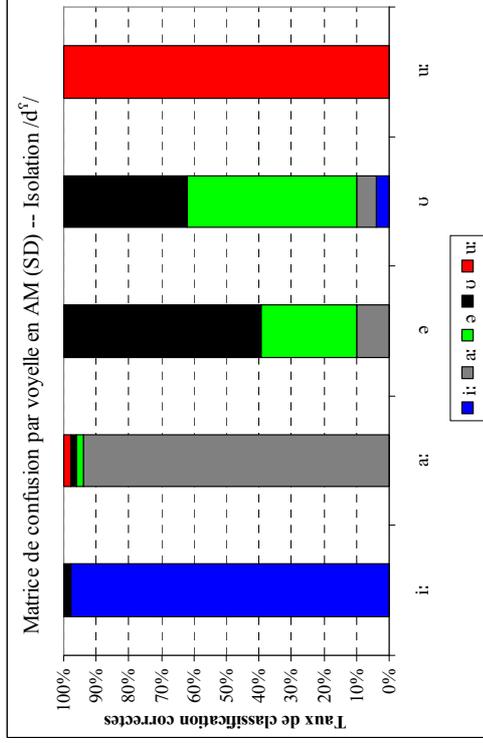
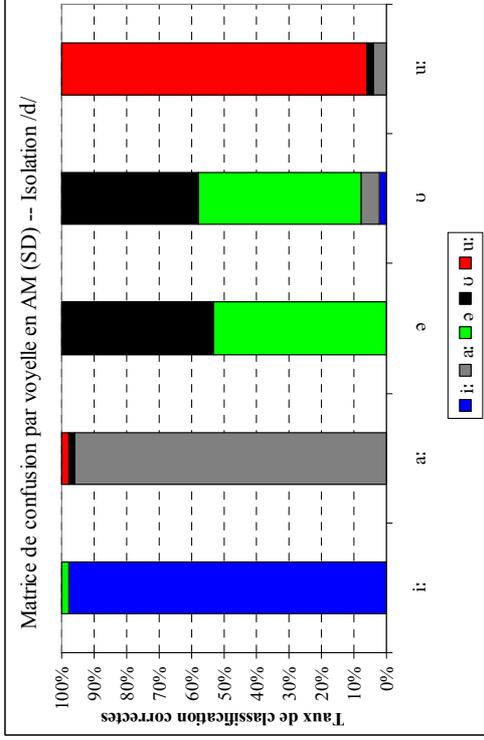
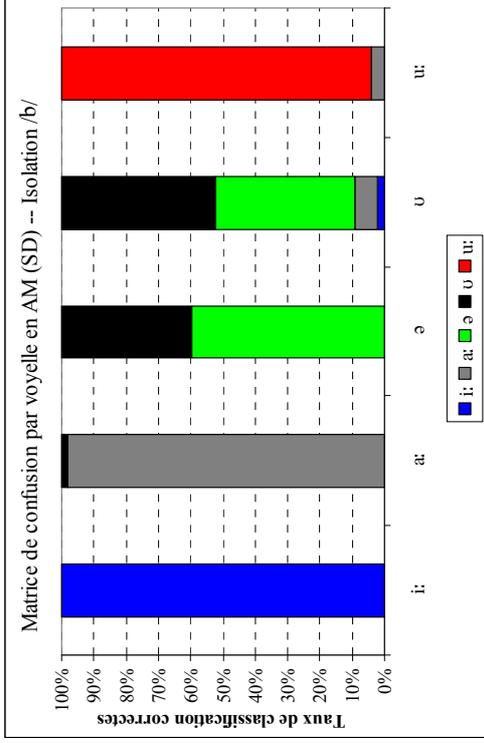
ANNEXE 15 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN STATIQUE AVEC LA DURÉE VOCALIQUE (AD) EN SITUATIONS SYLLABE ET ISOLATION.



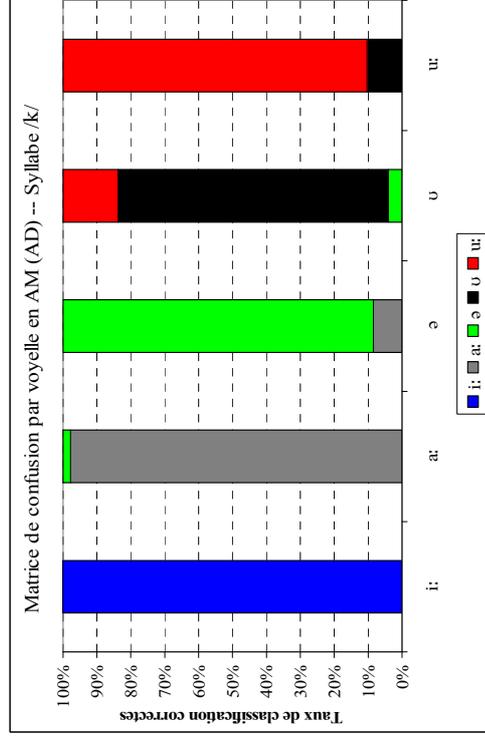
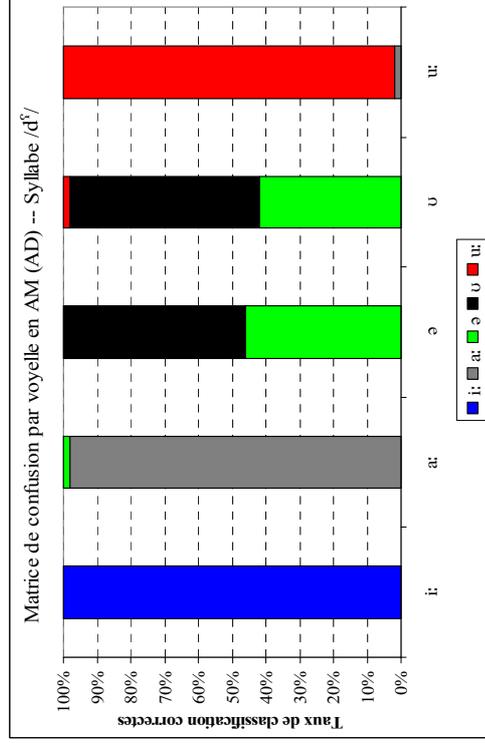
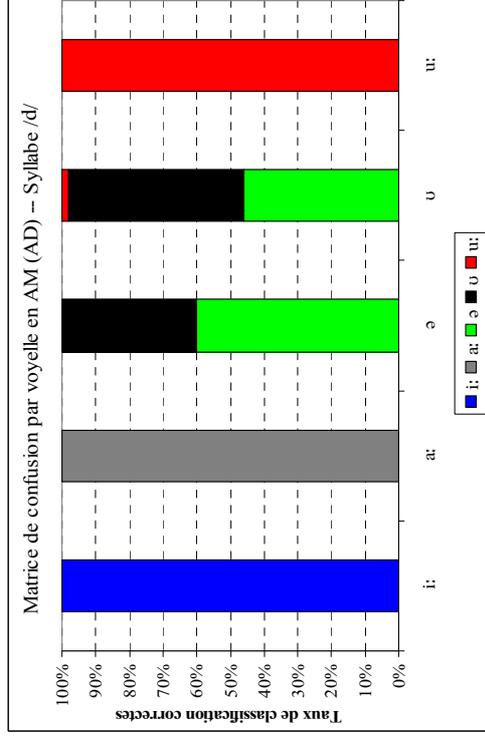
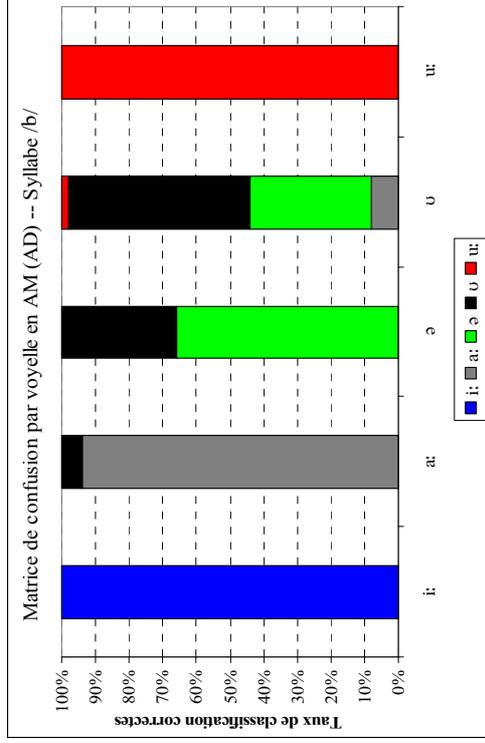


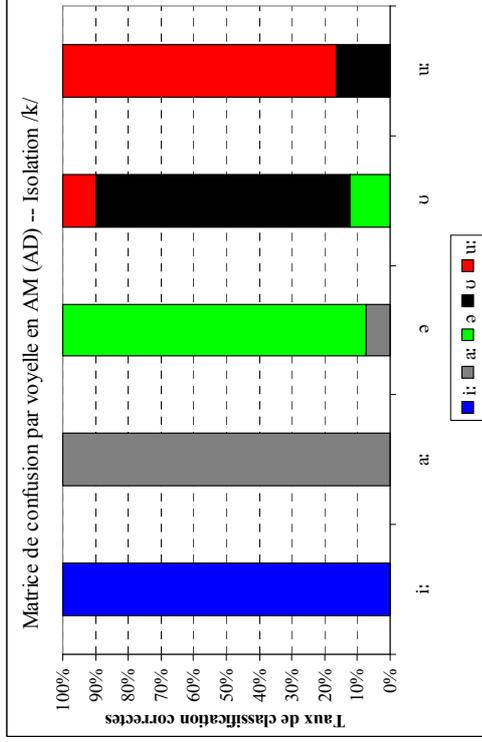
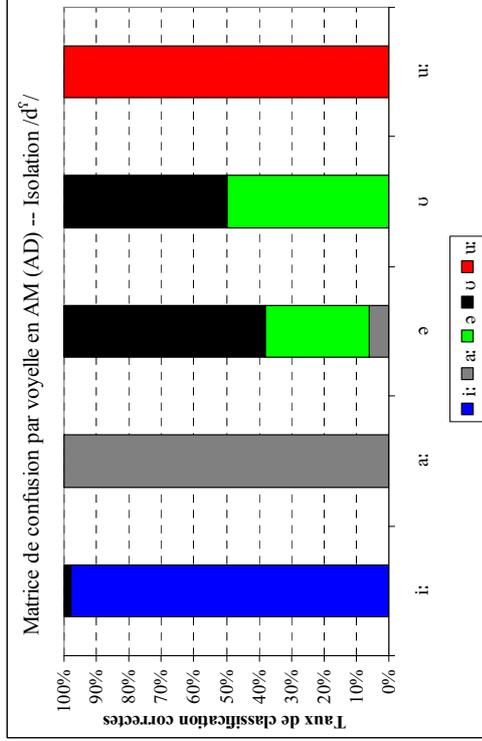
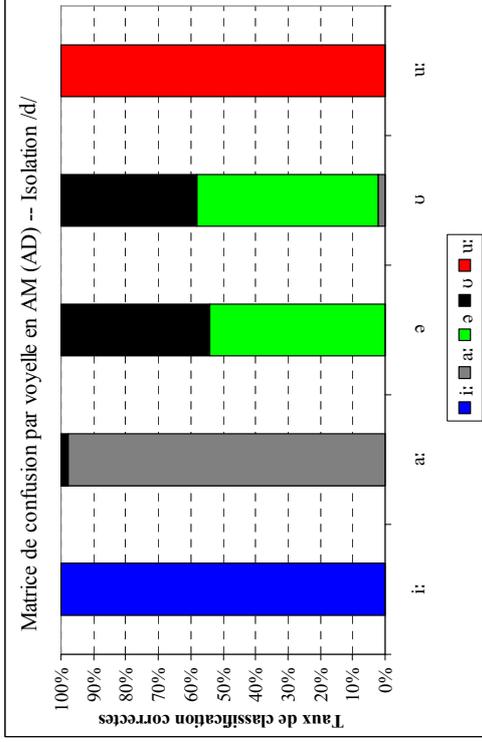
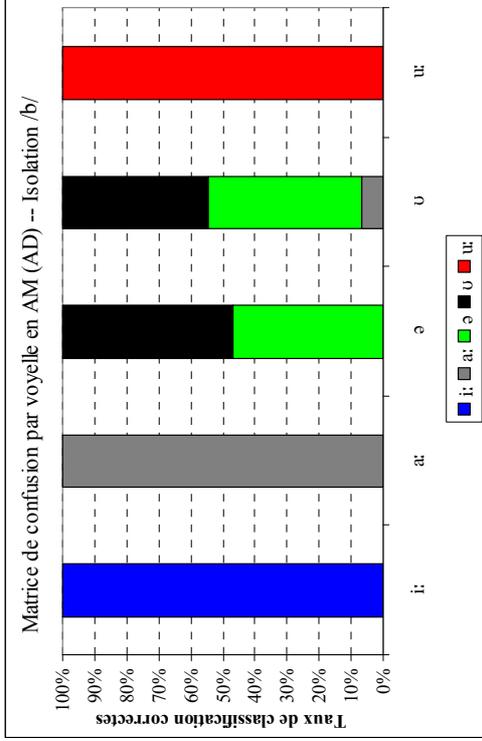
ANNEXE 16 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN STATIQUE SANS LA DURÉE VOCALIQUE (SD) EN SITUATIONS SYLLABE ET ISOLATION.



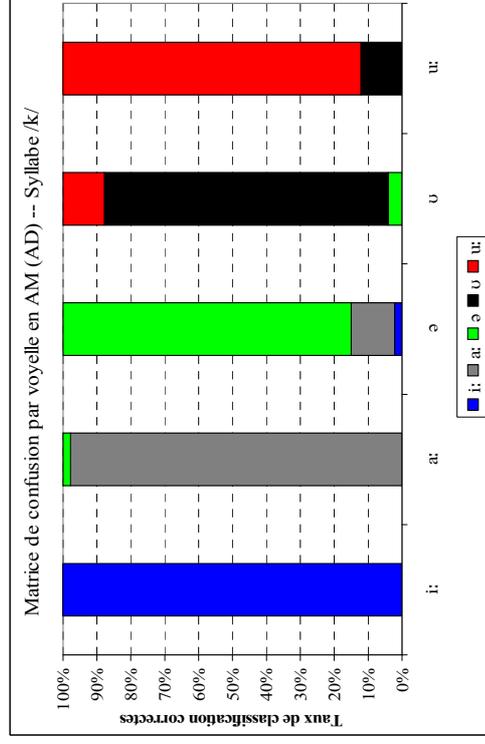
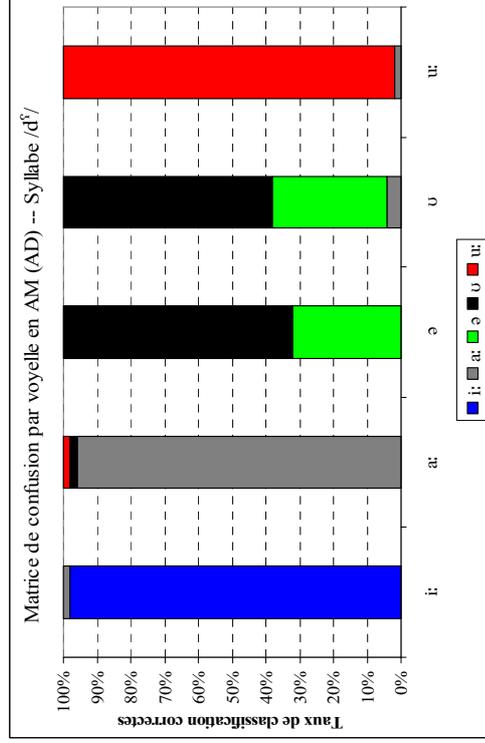
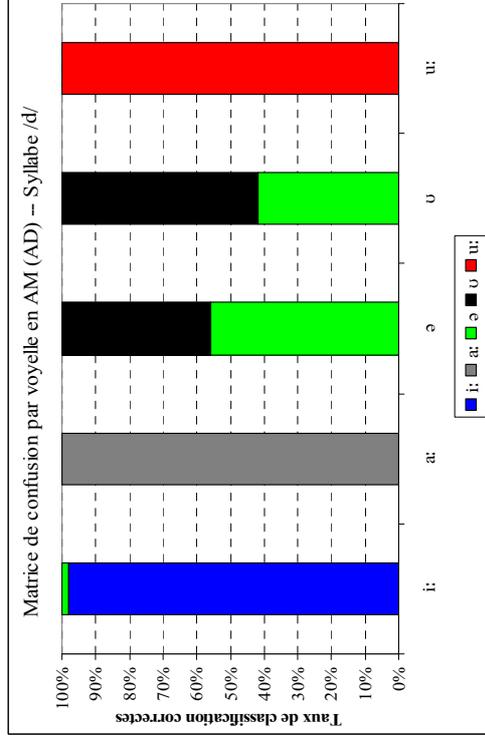
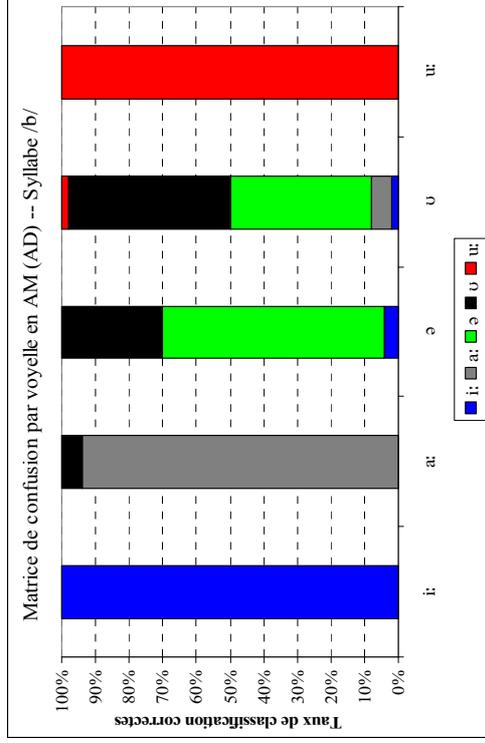


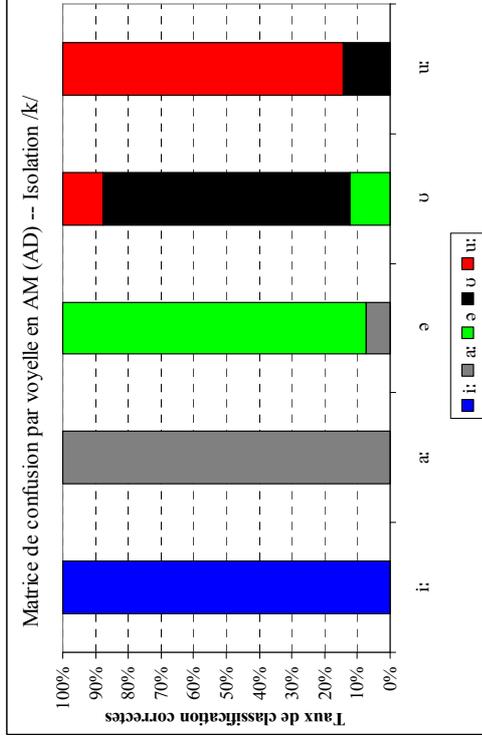
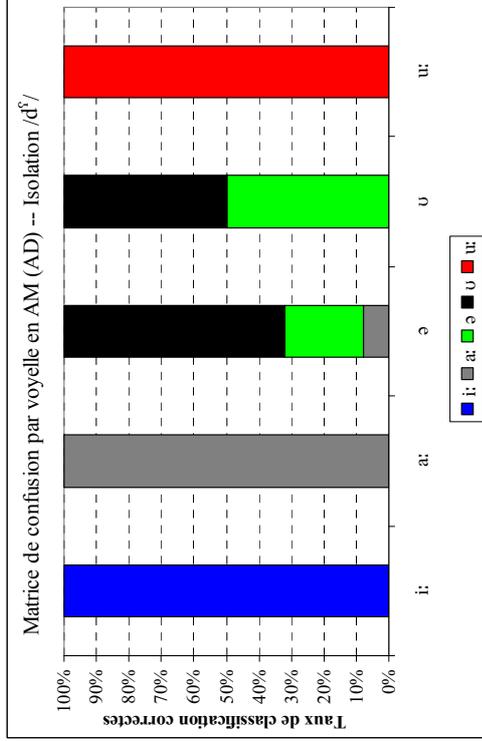
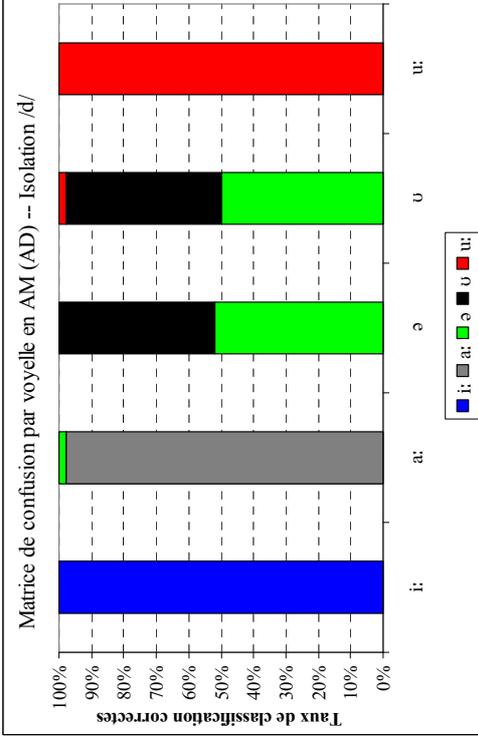
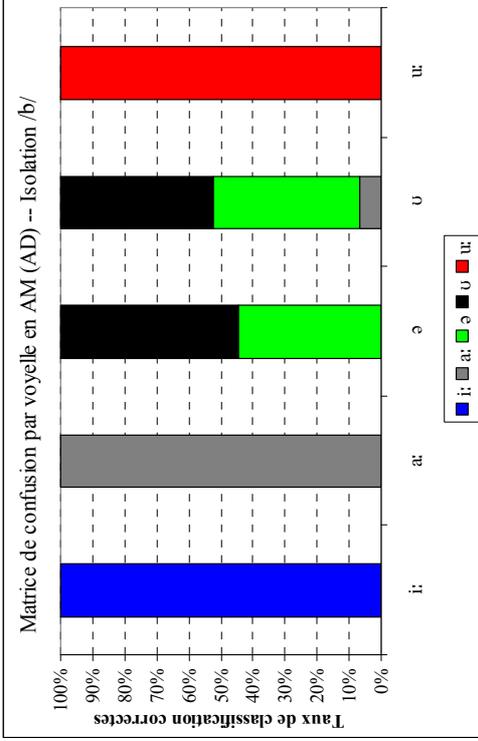
ANNEXE 17 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.





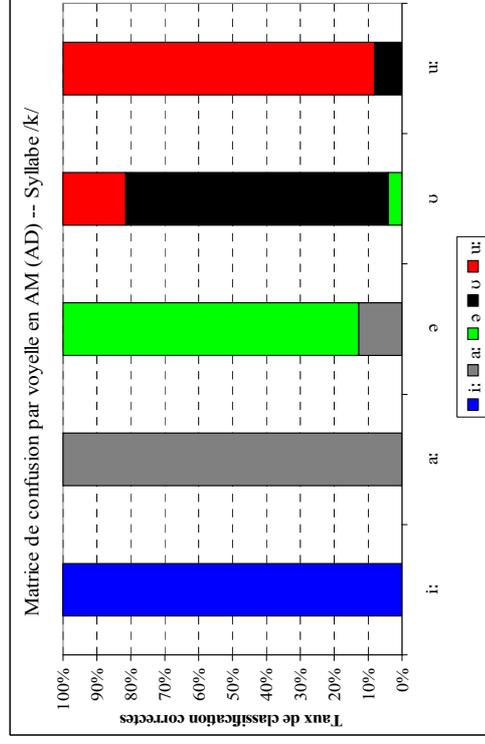
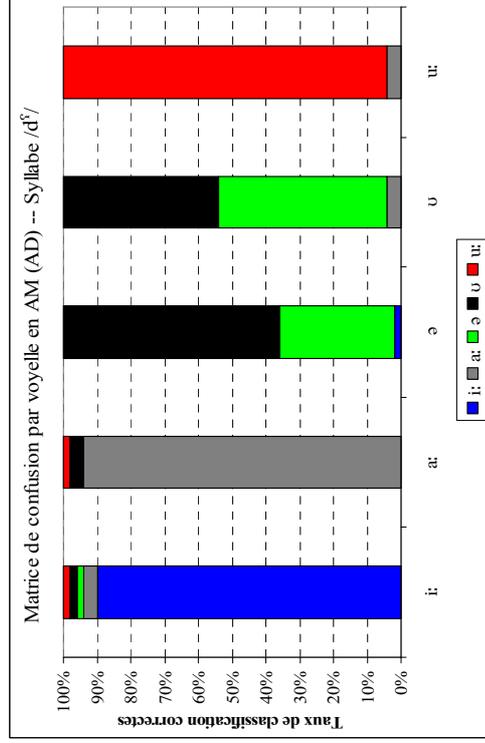
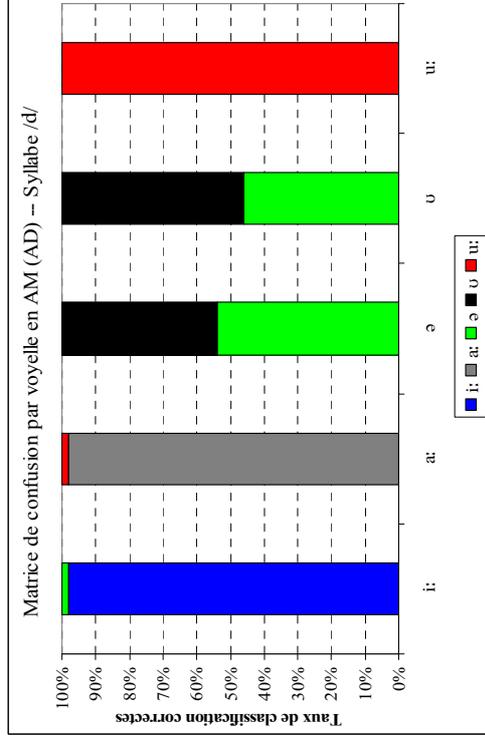
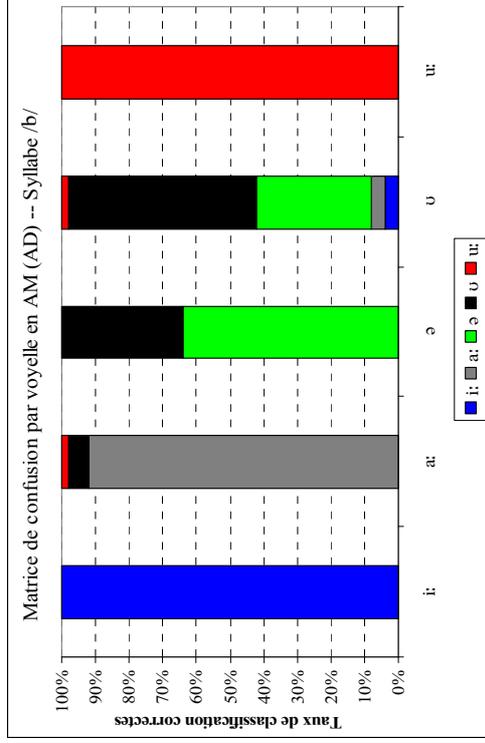
ANNEXE 18 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.

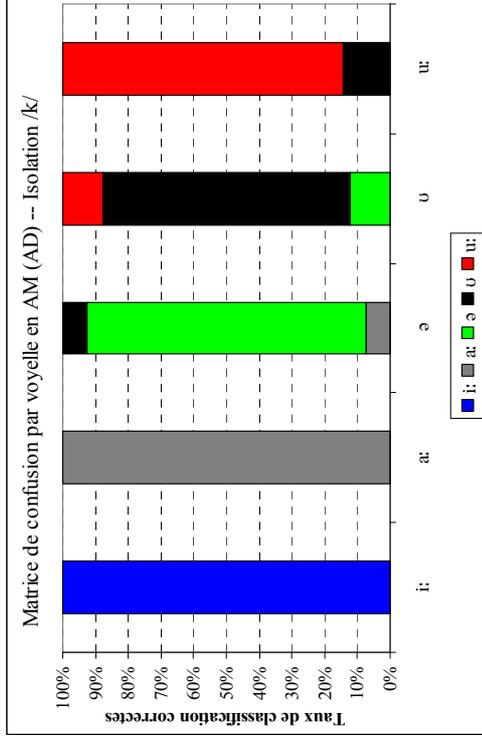
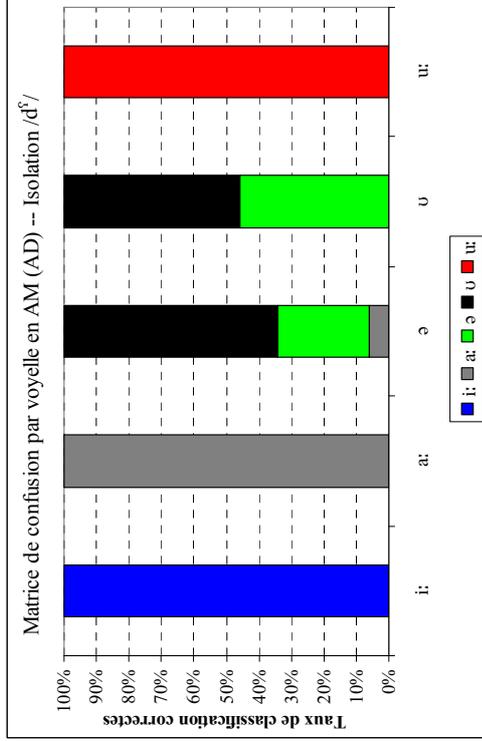
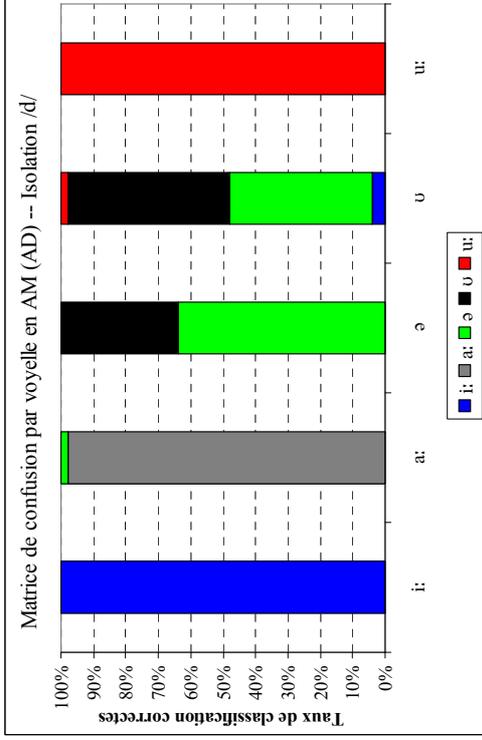
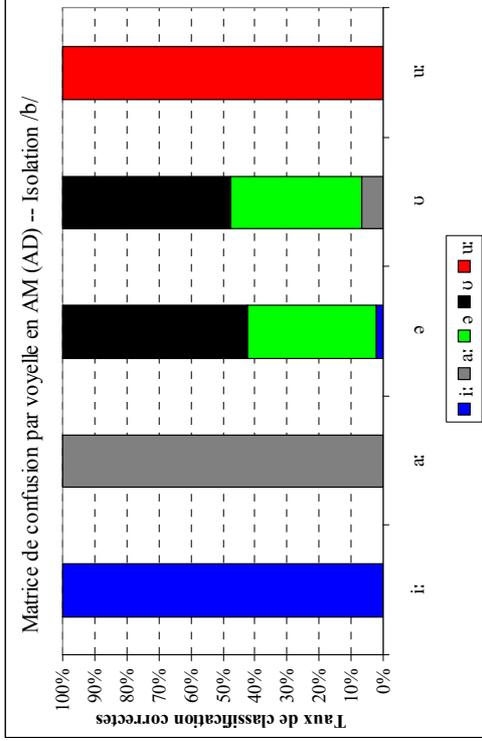




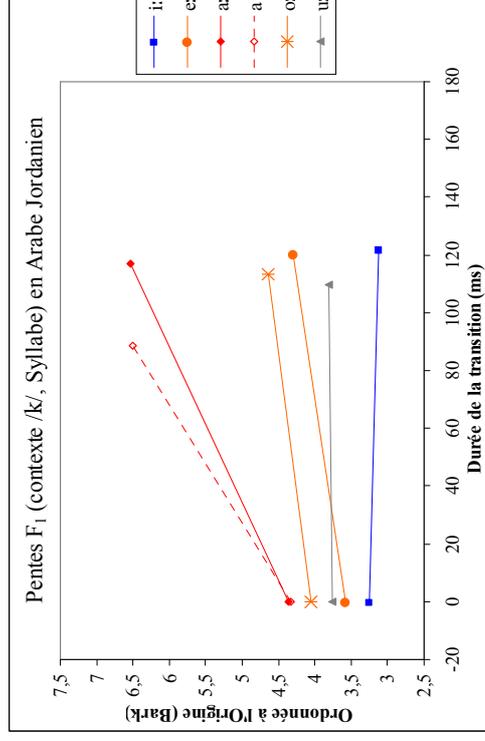
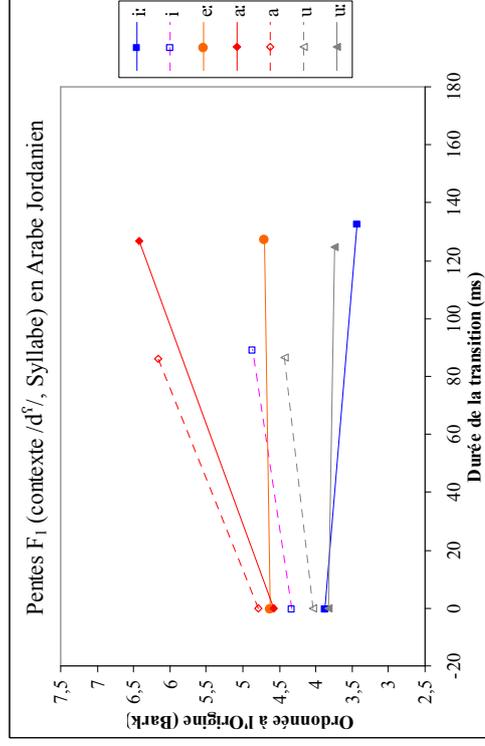
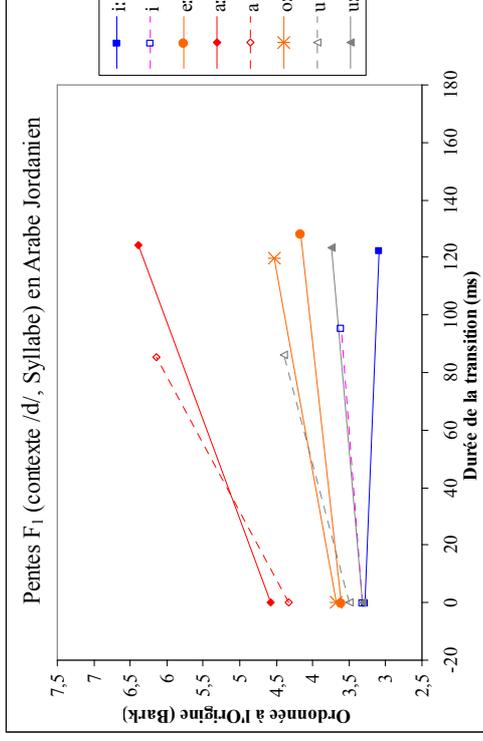
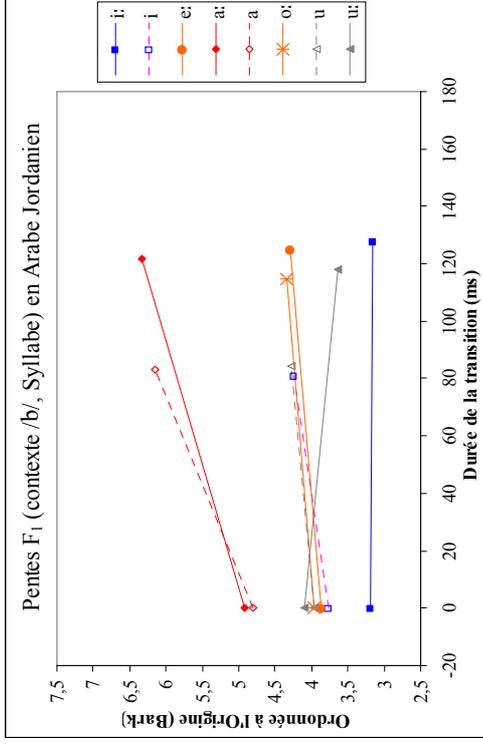
ANNEXE 19

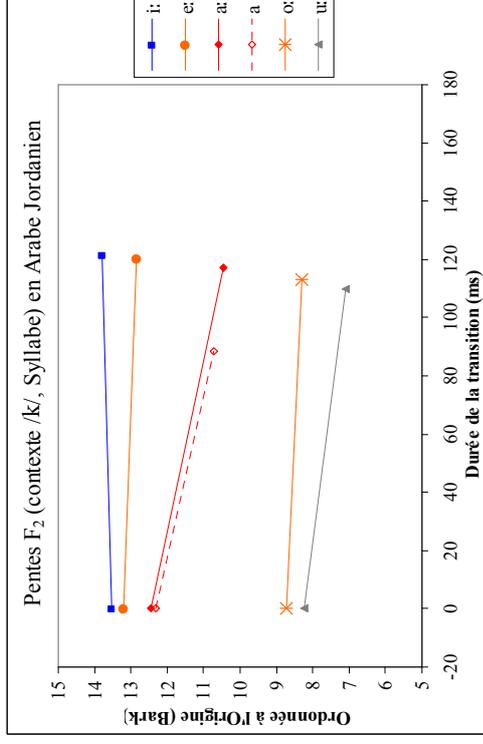
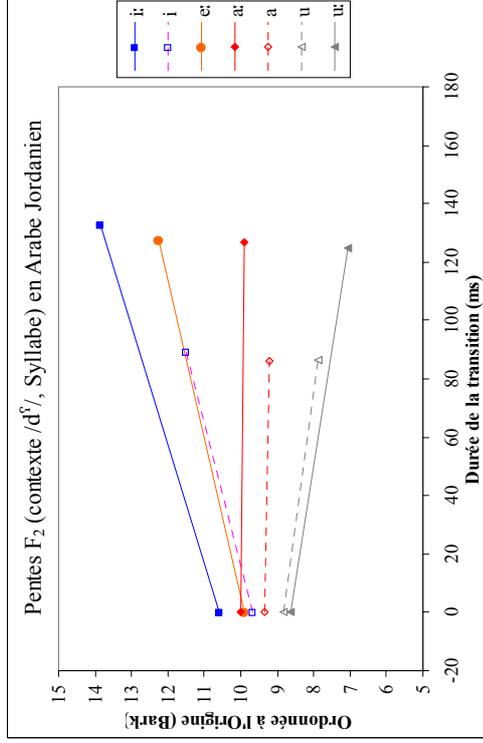
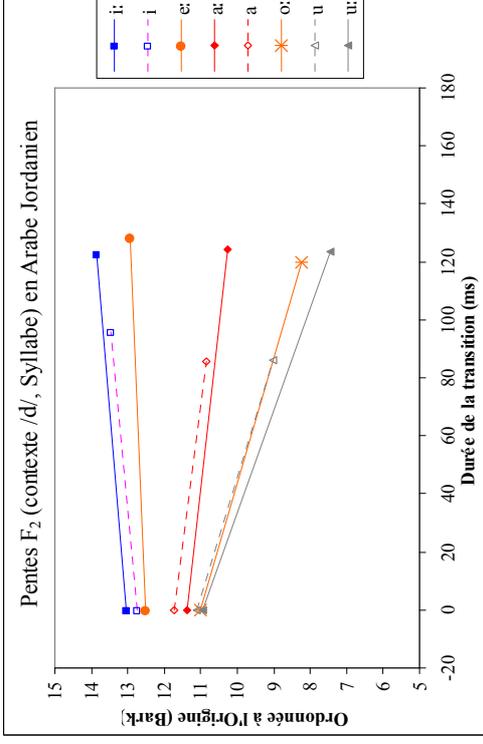
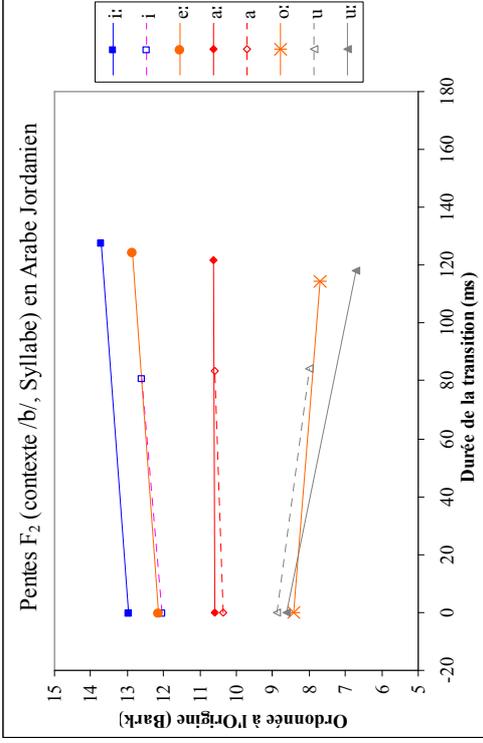
MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE MAROCAIN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.

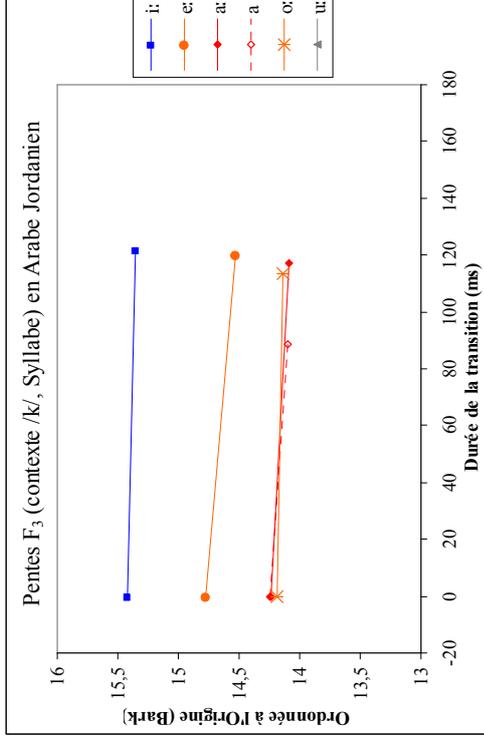
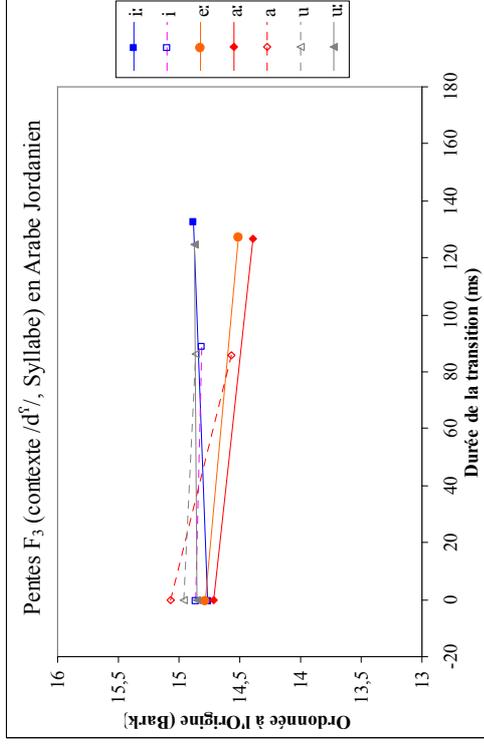
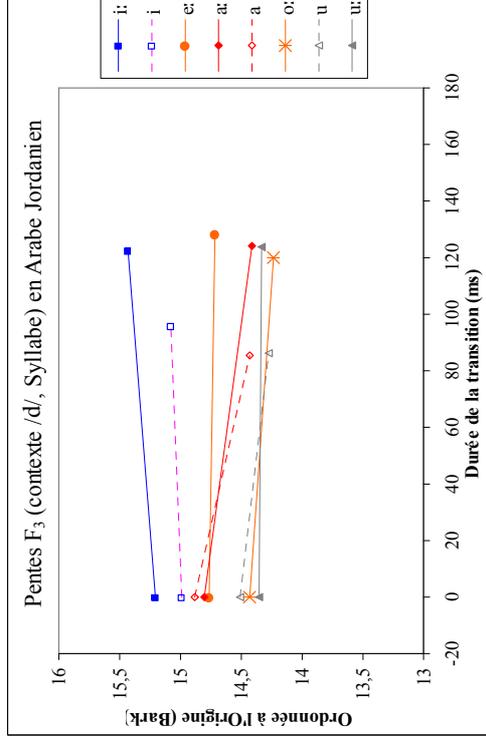
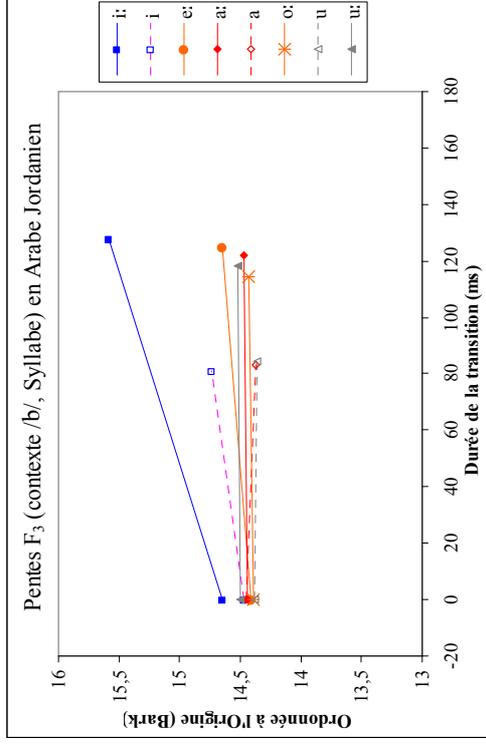


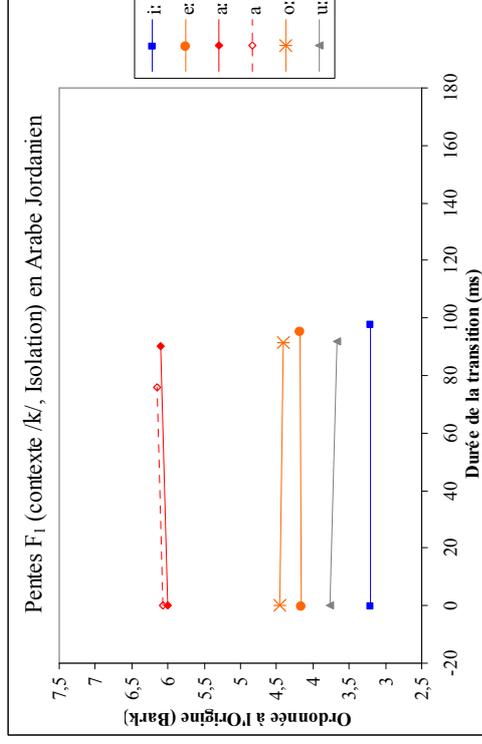
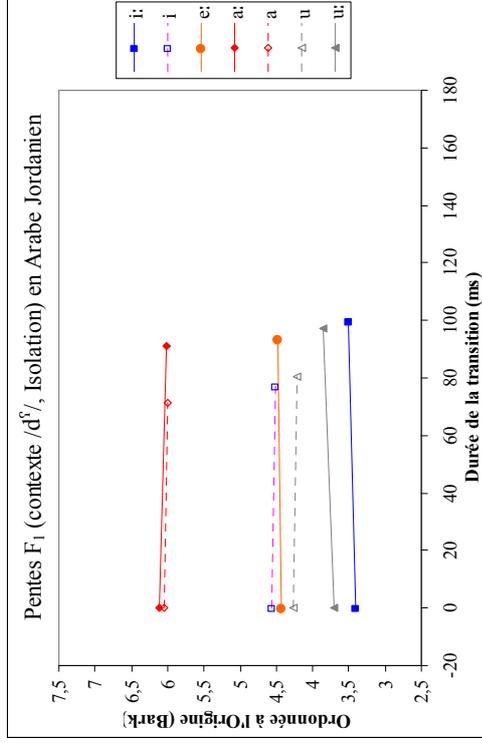
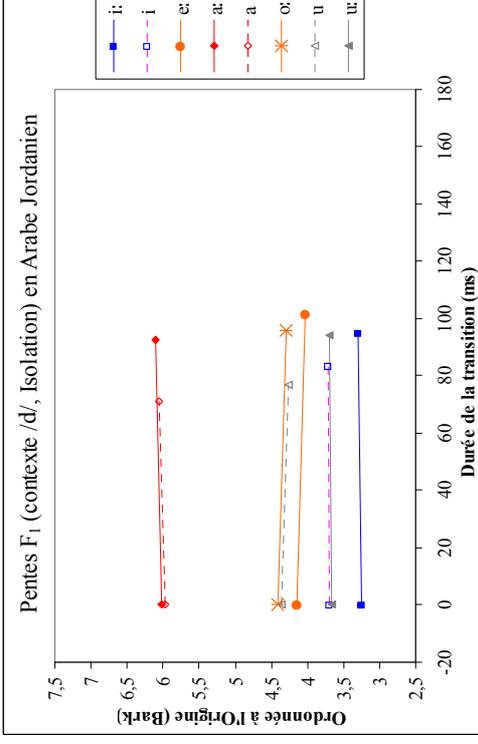
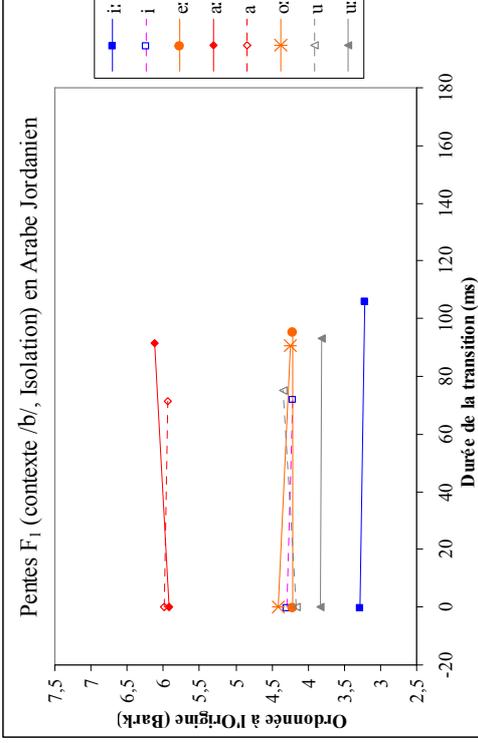


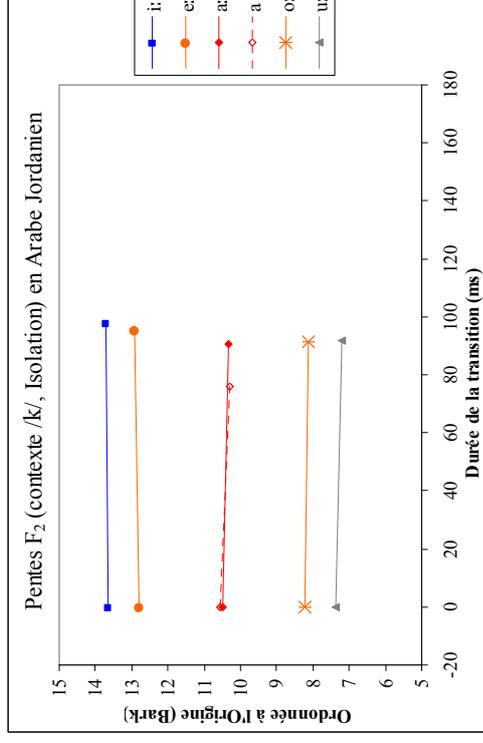
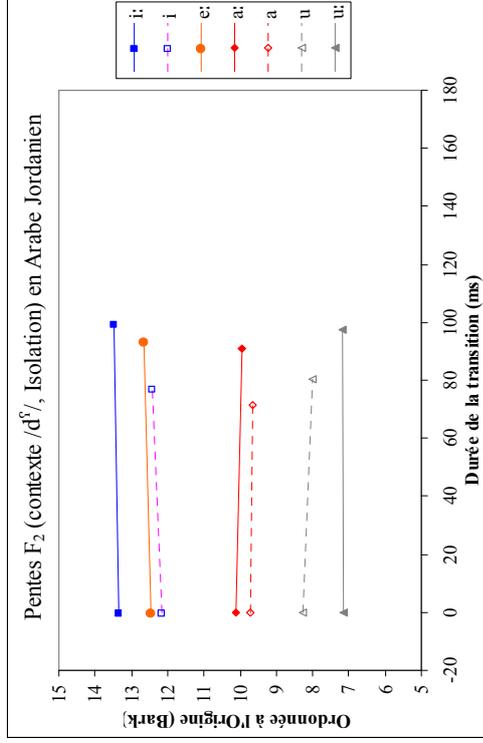
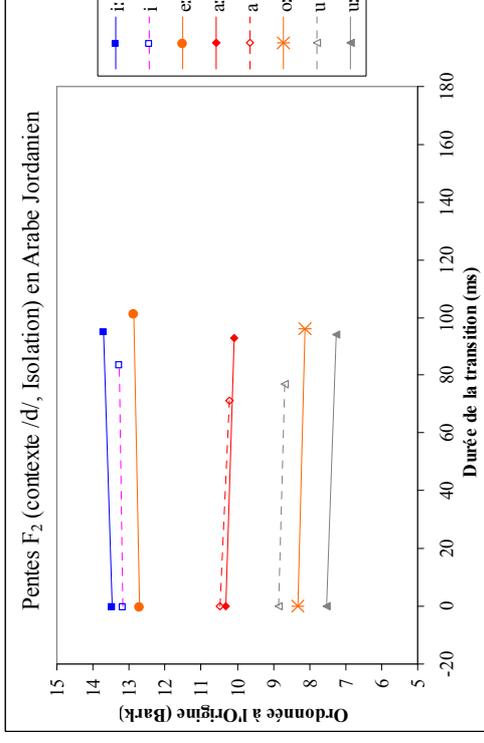
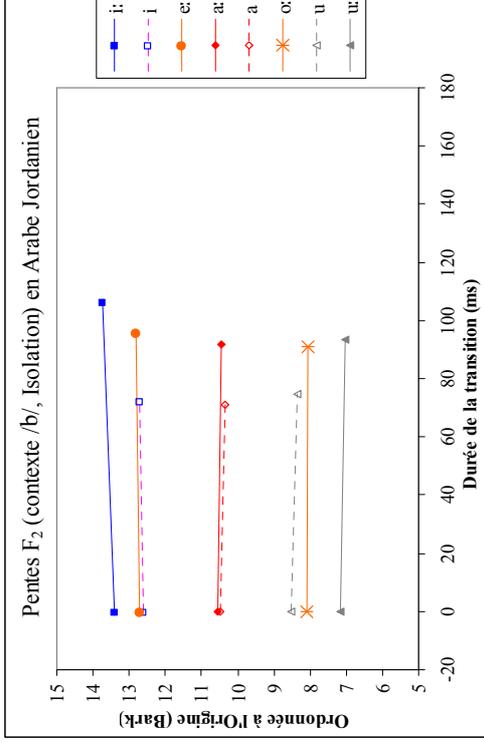
ANNEXE 20 PENTES FORMANTIQUES (F_1 , F_2 , F_3) PAR VOYELLE EN SYLLABE & ISOLATION EN ARABE JORDANIEN

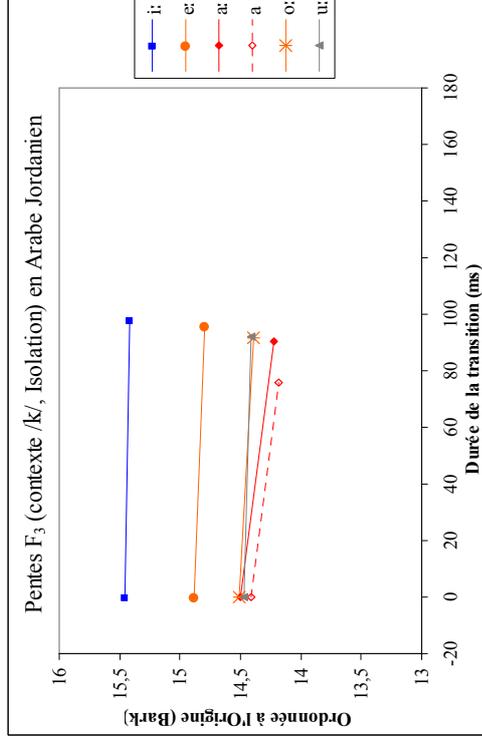
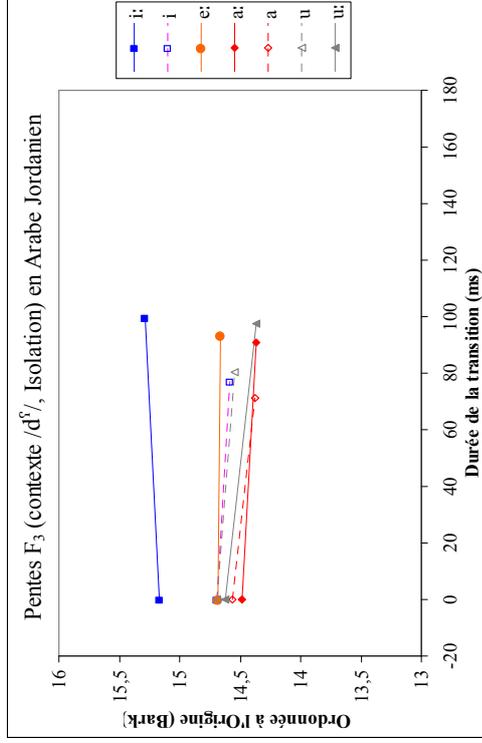
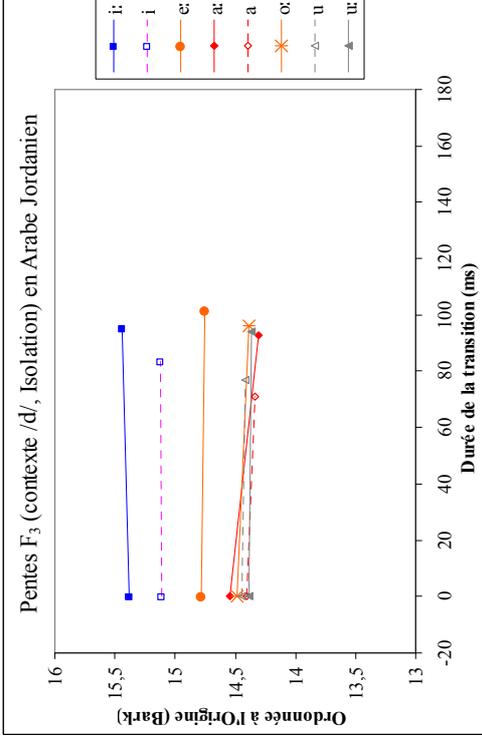
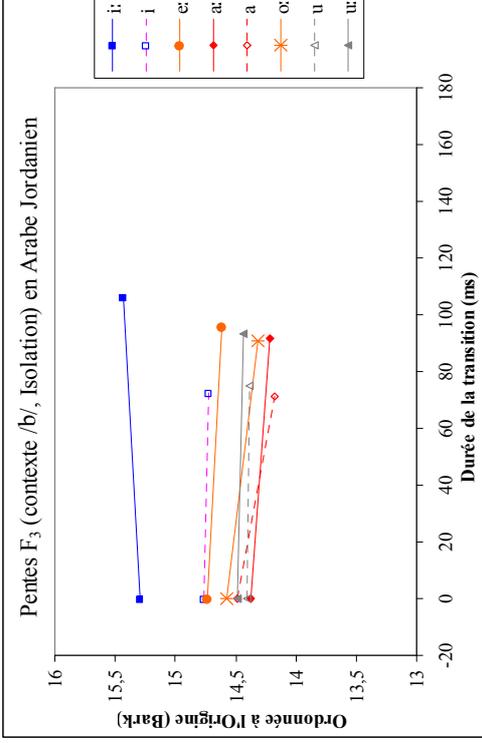




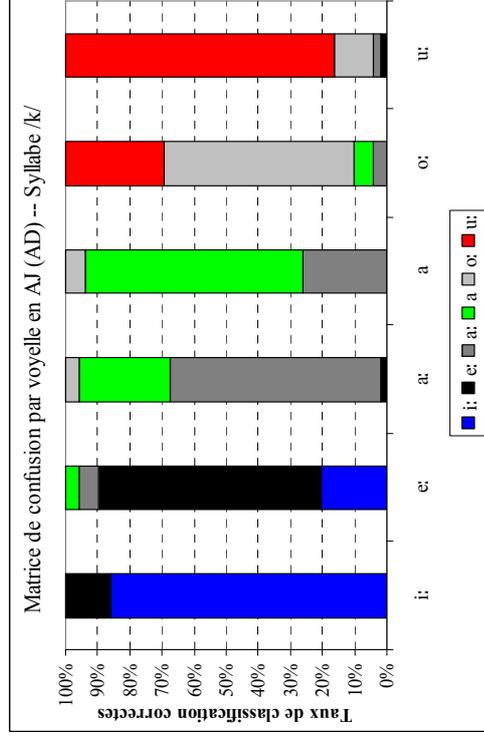
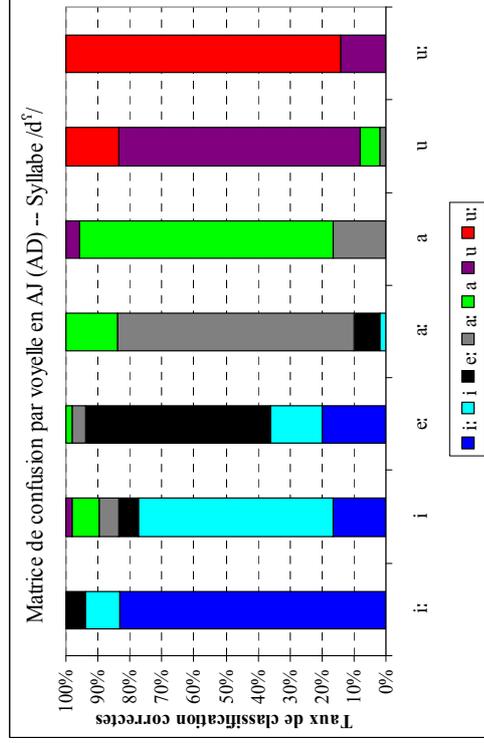
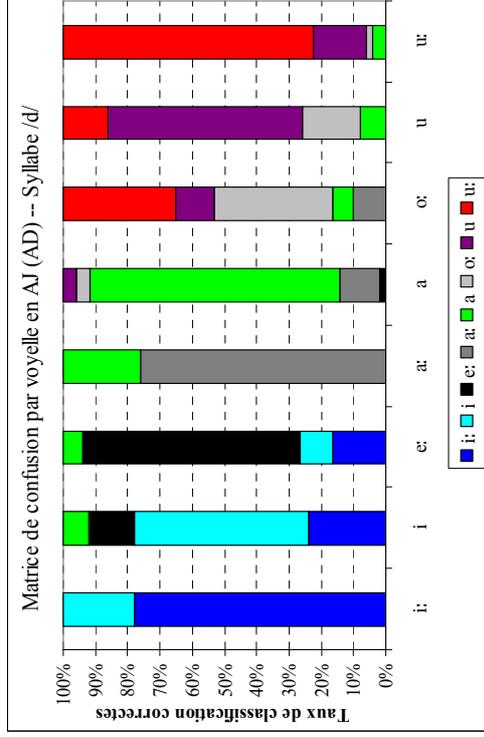
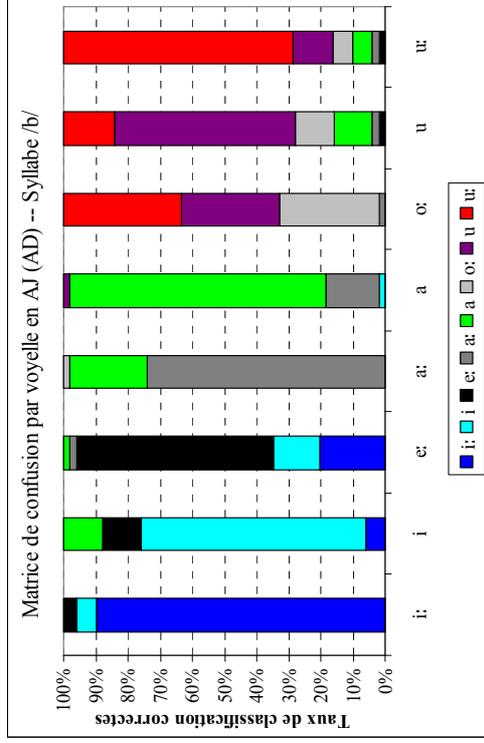


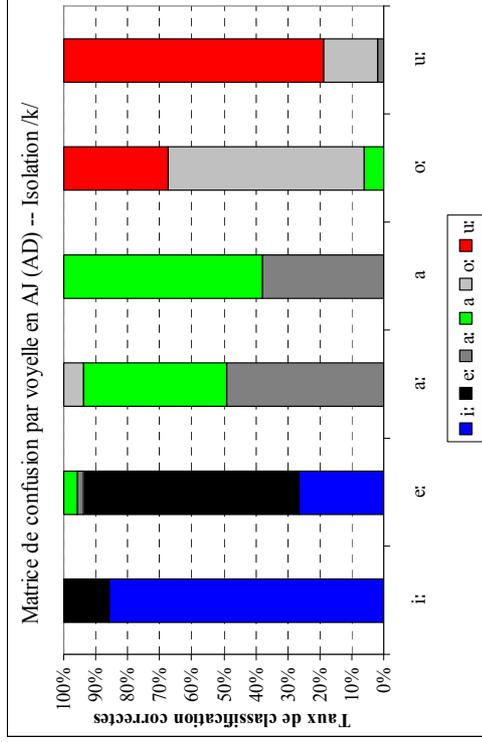
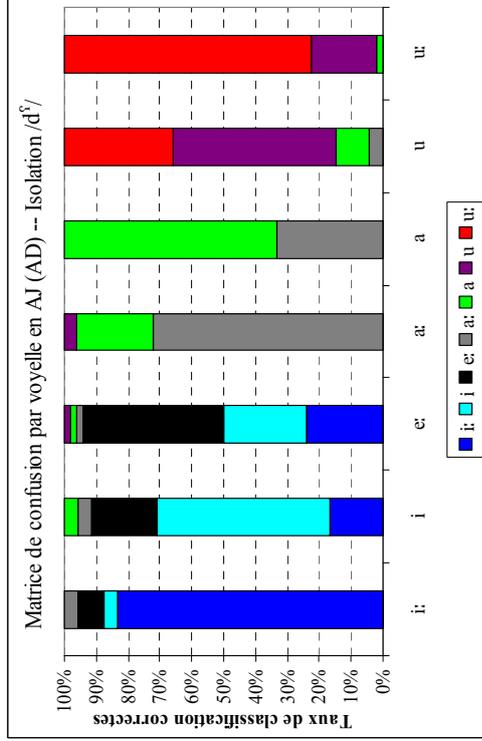
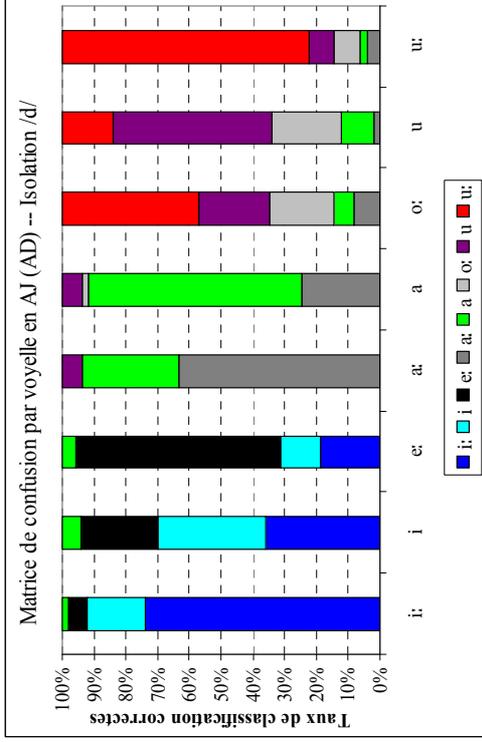
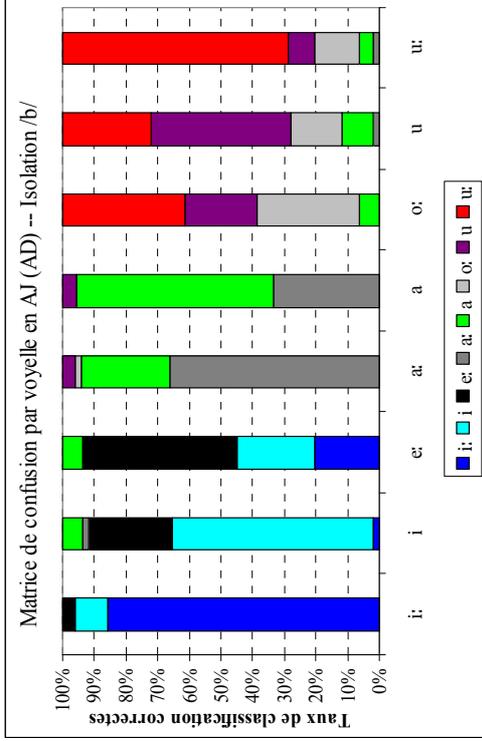




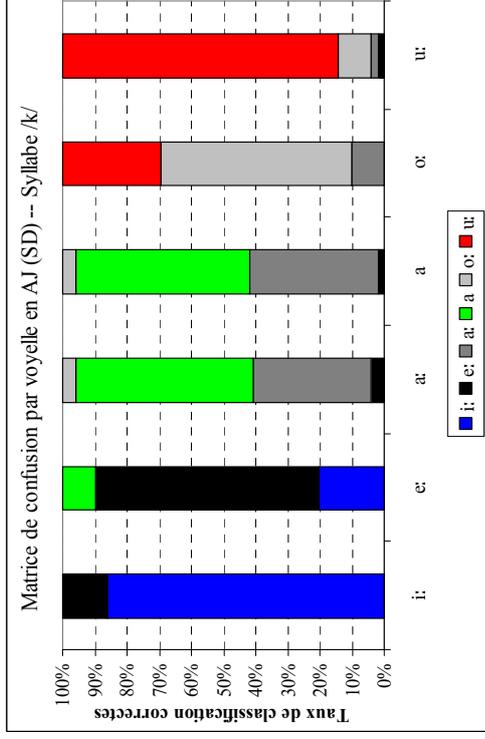
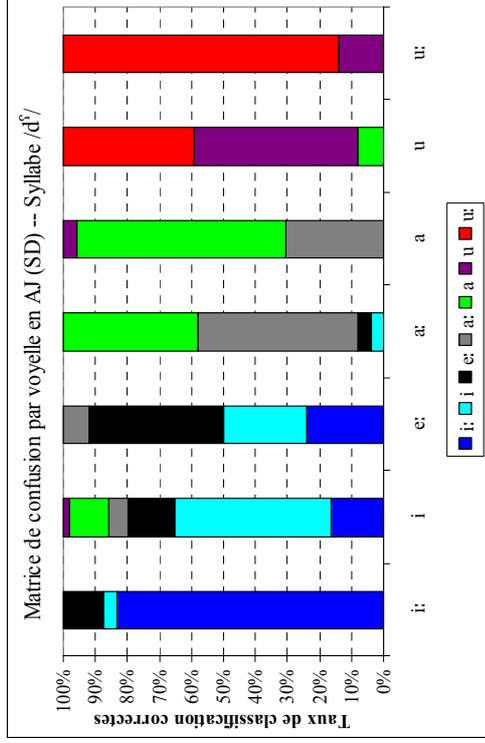
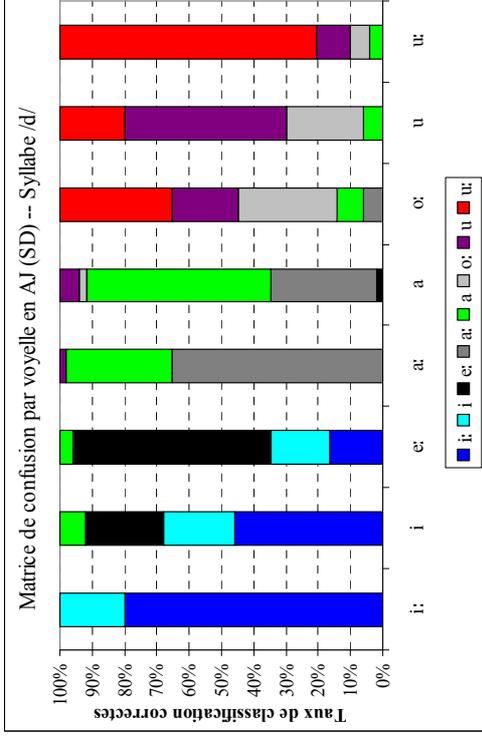
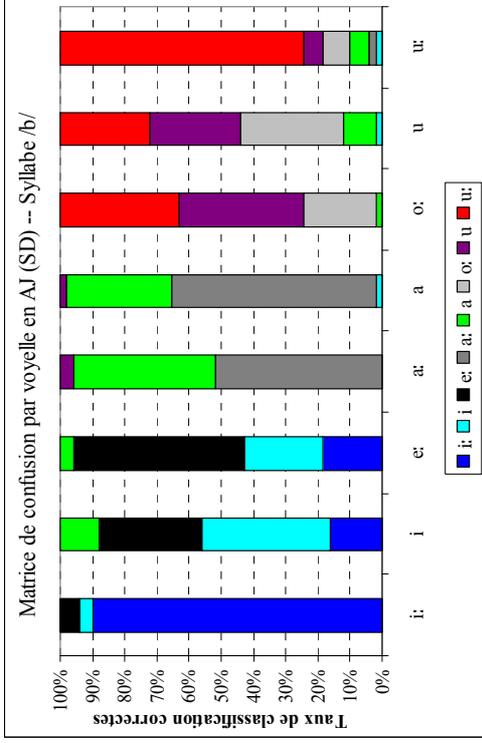


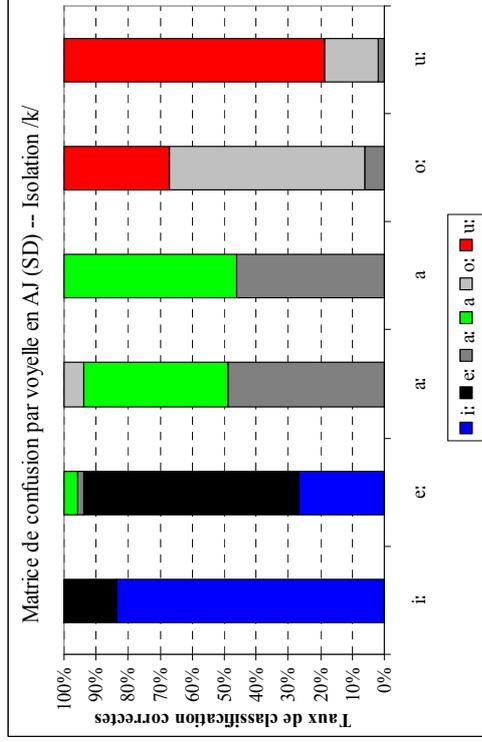
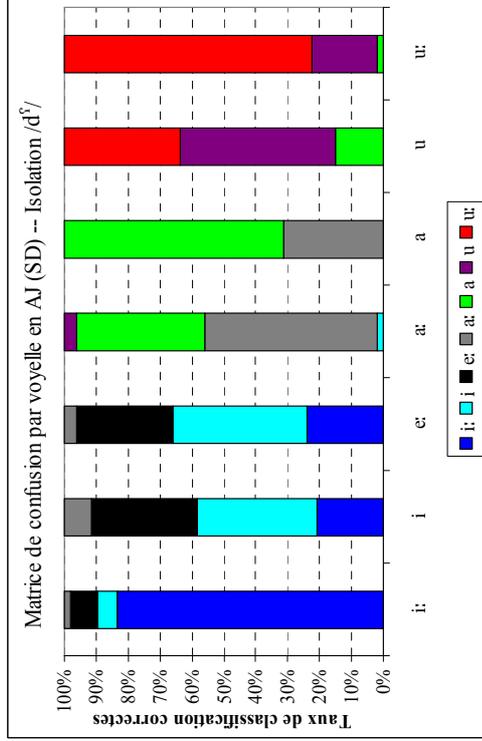
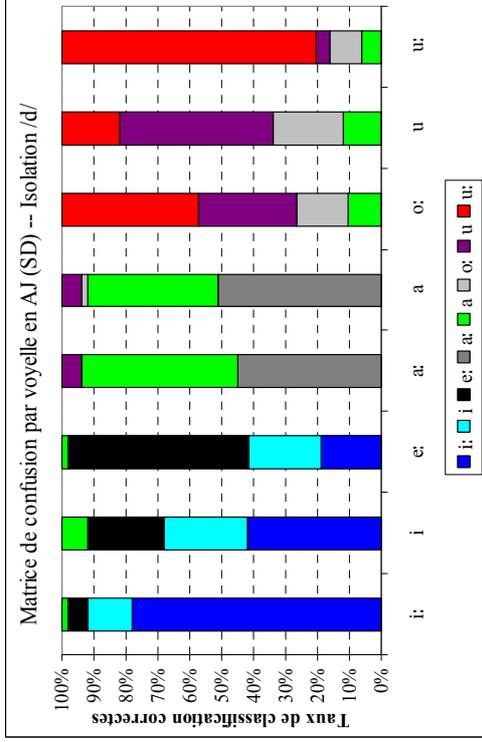
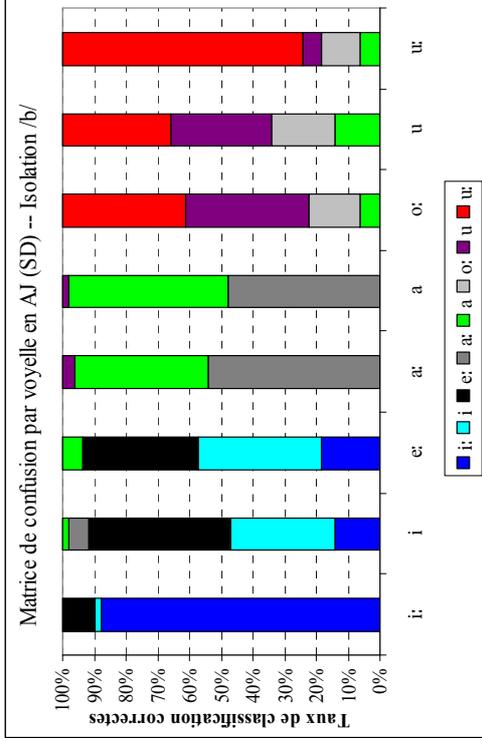
ANNEXE 21 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN STATIQUE AVEC LA DURÉE VOCALIQUE (AD) EN SITUATIONS SYLLABE ET ISOLATION.





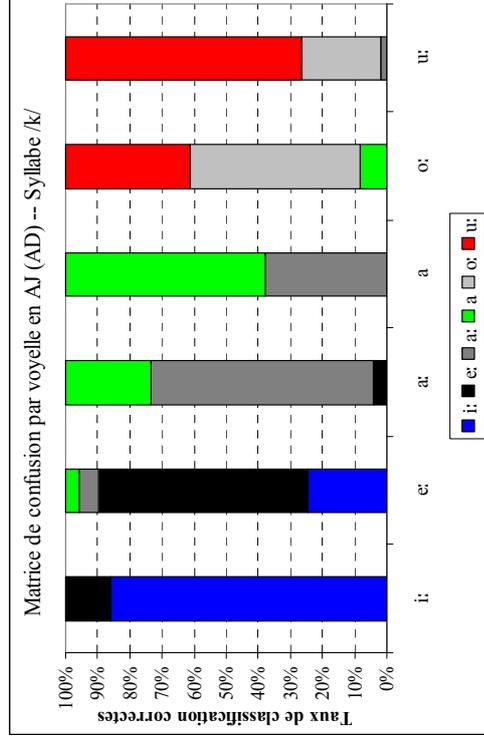
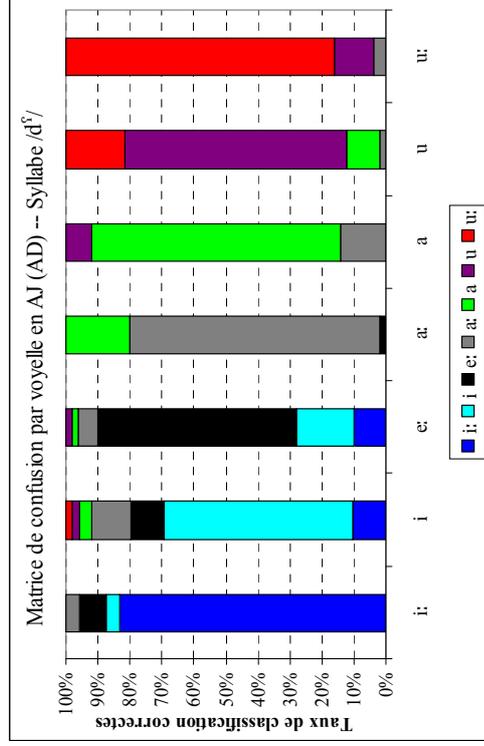
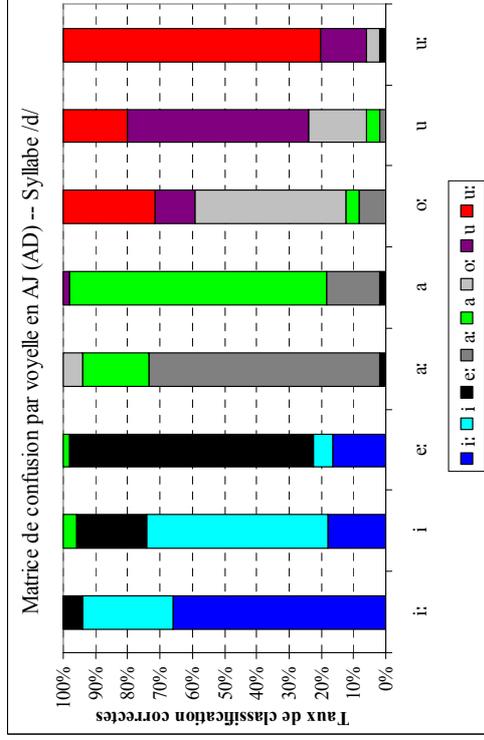
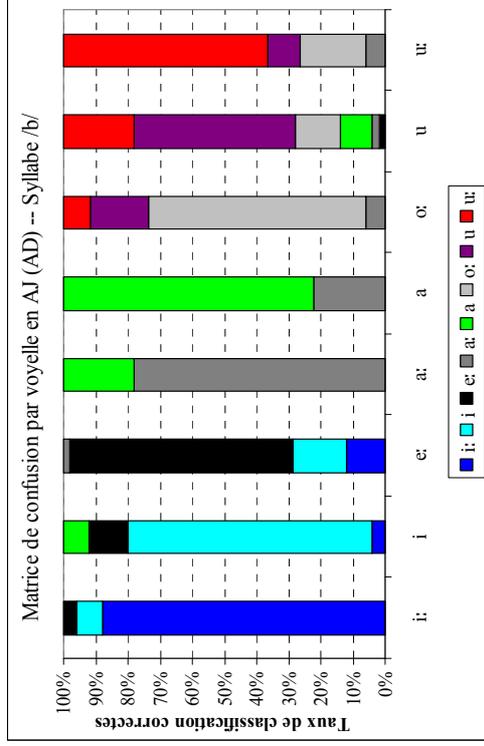
ANNEXE 22 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN STATIQUE SANS LA DURÉE VOCALIQUE (SD) EN SITUATIONS SYLLABE ET ISOLATION.

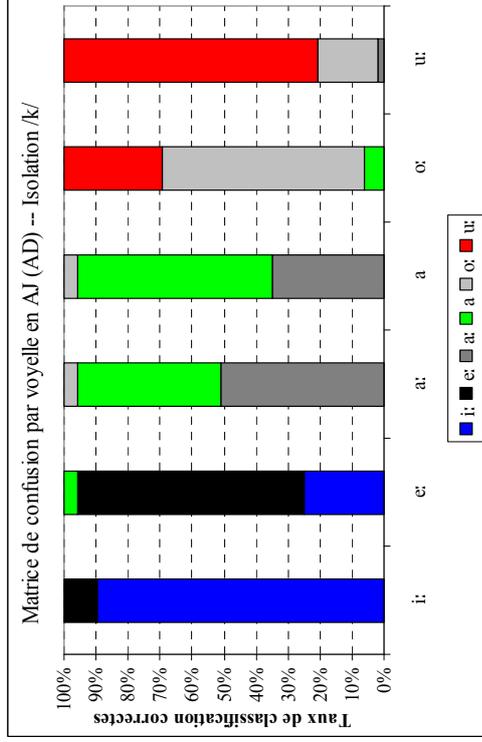
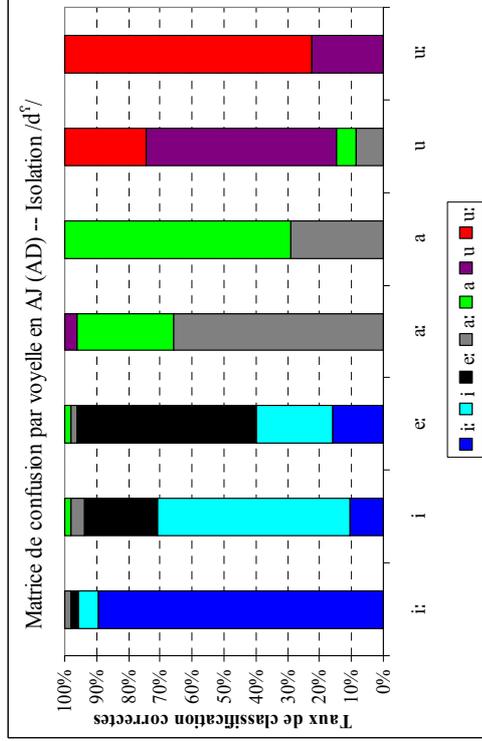
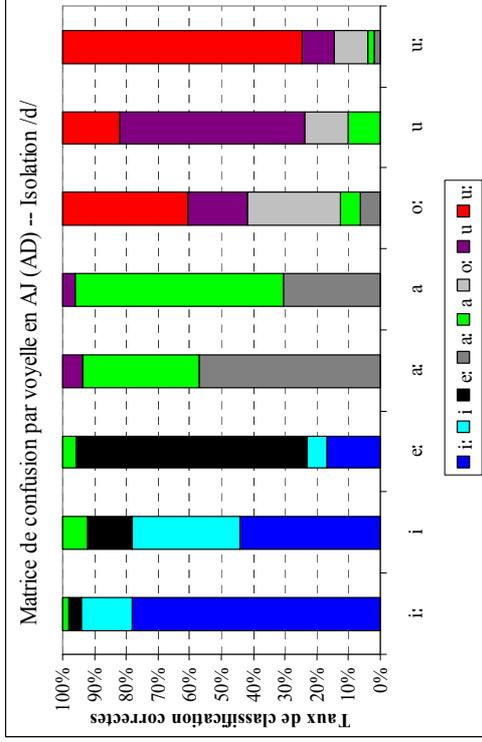
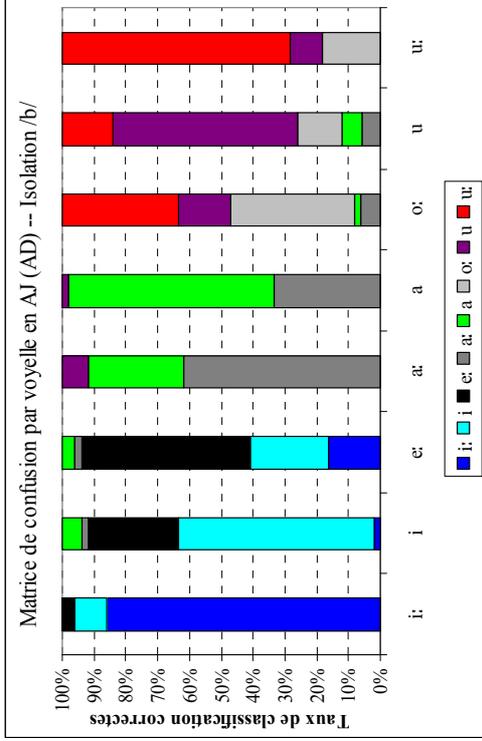




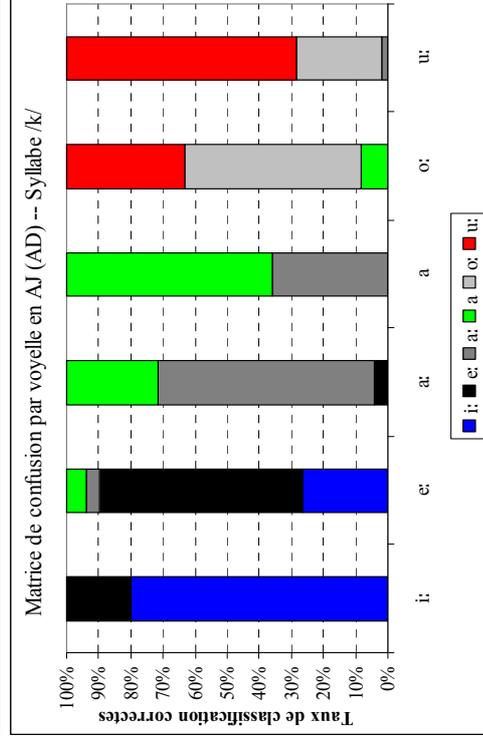
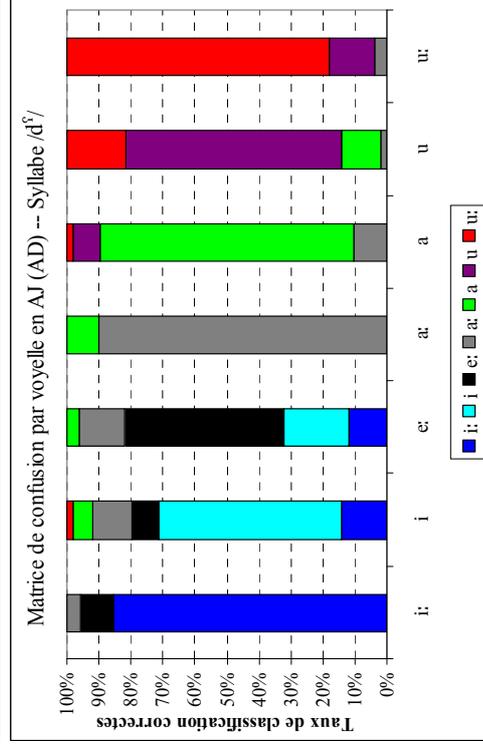
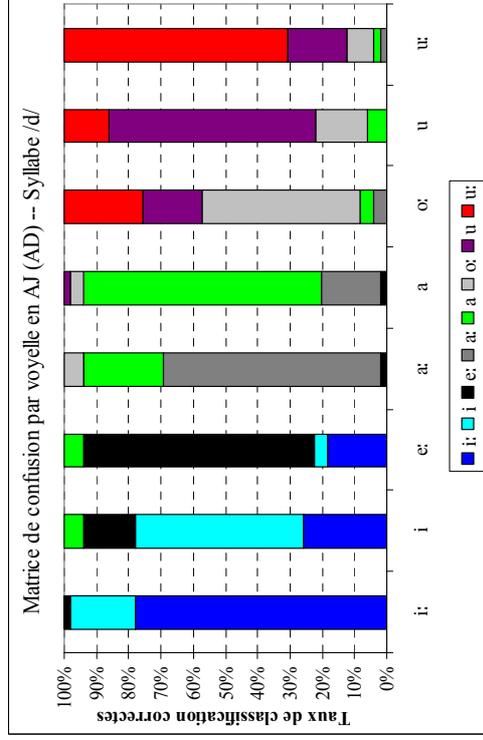
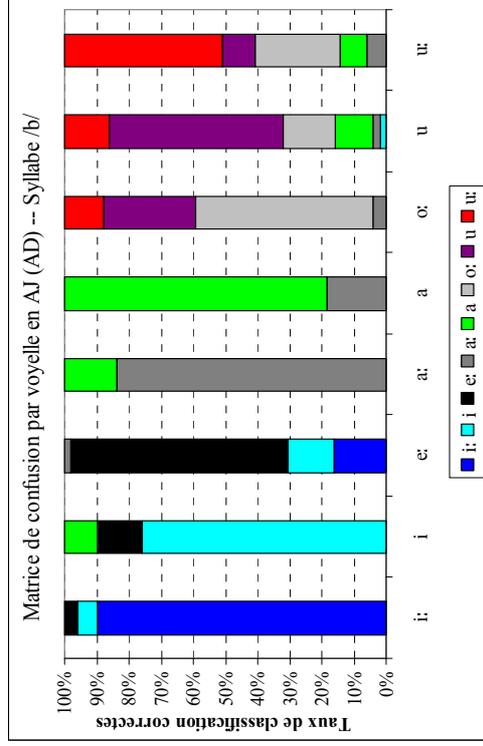
ANNEXE 23

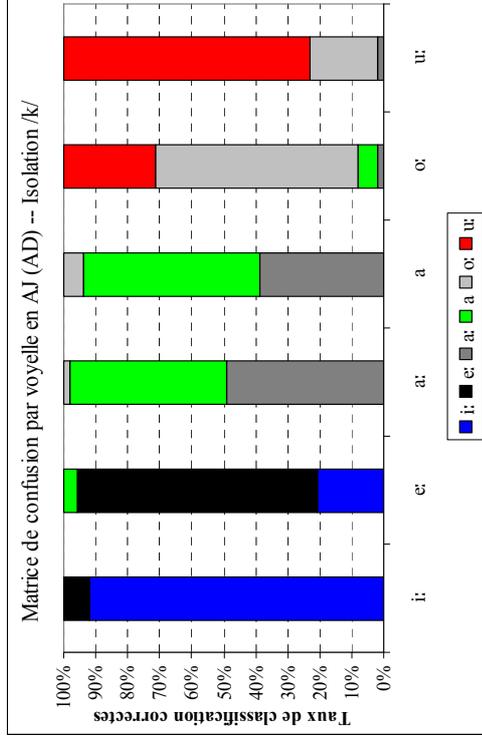
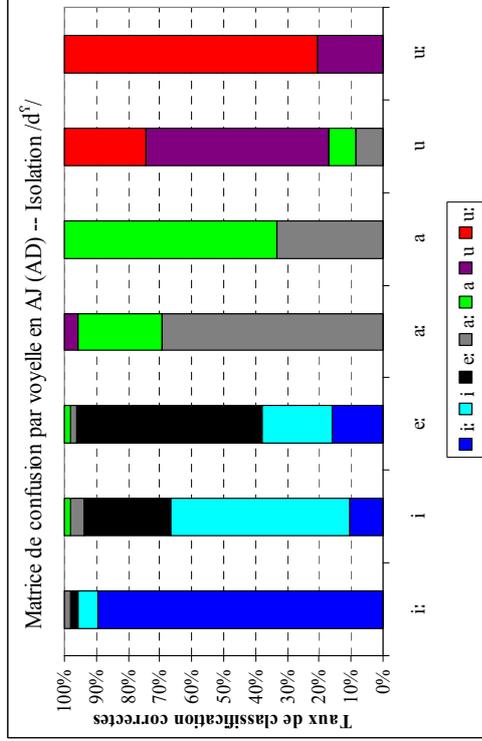
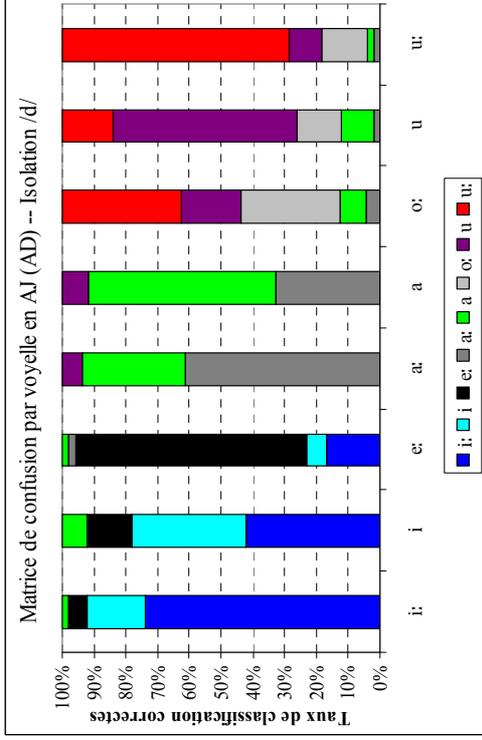
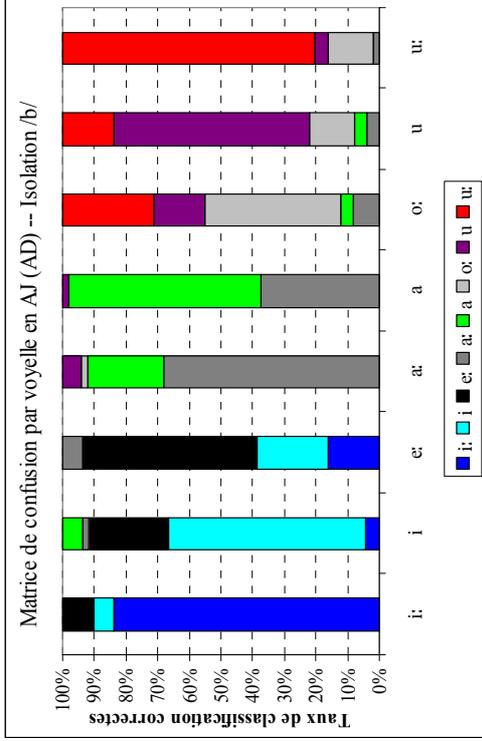
MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.





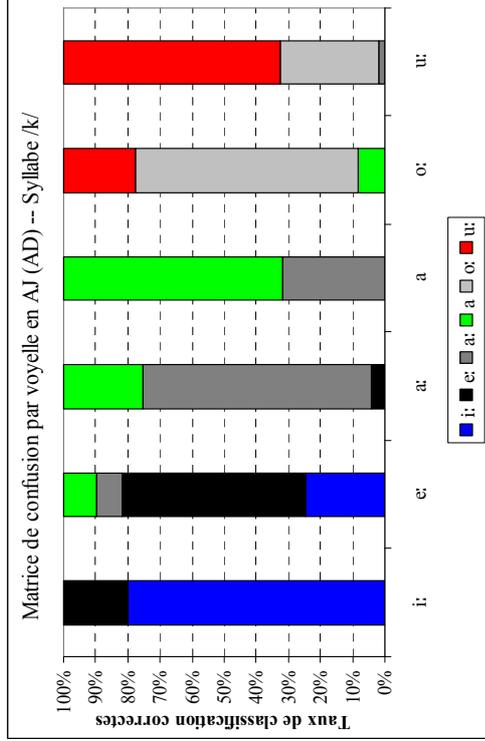
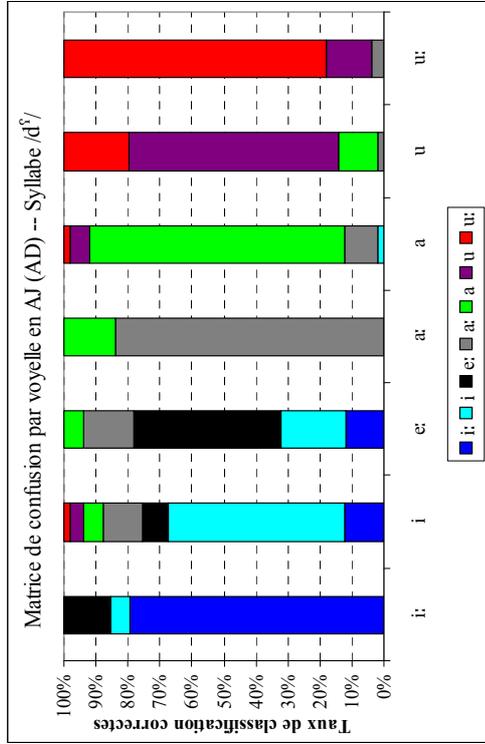
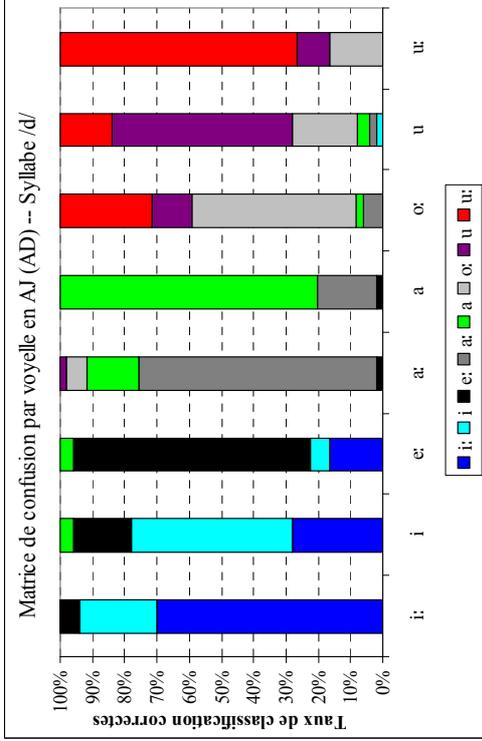
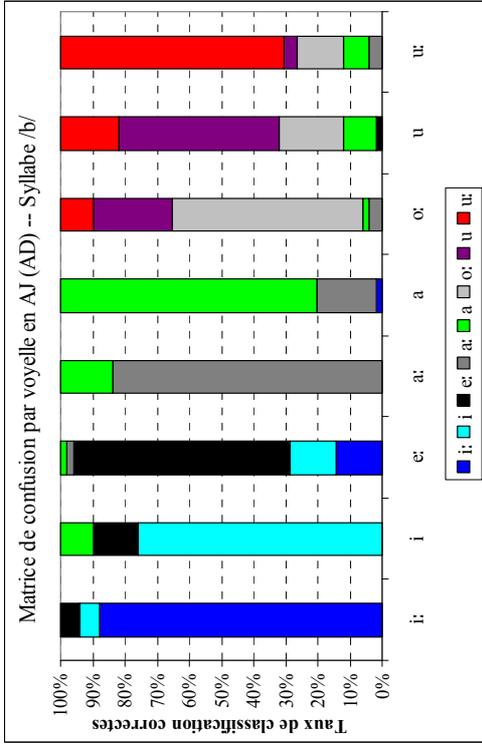
ANNEXE 24 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.

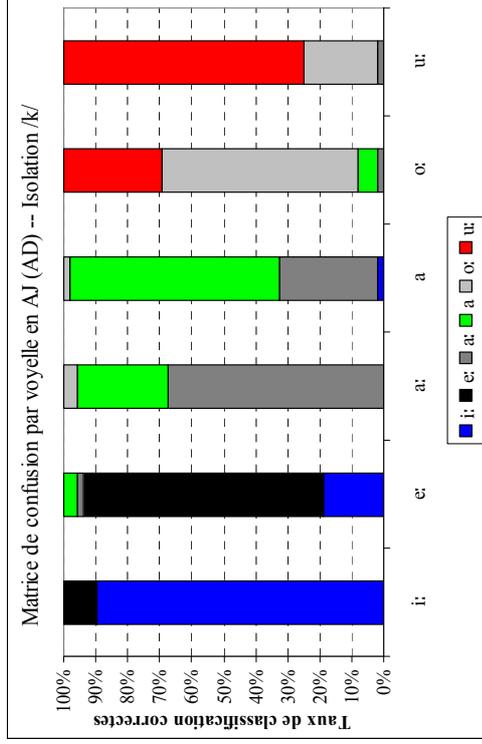
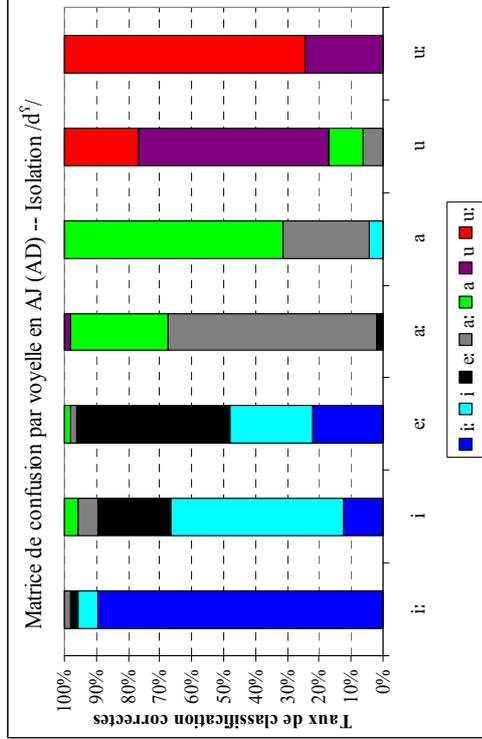
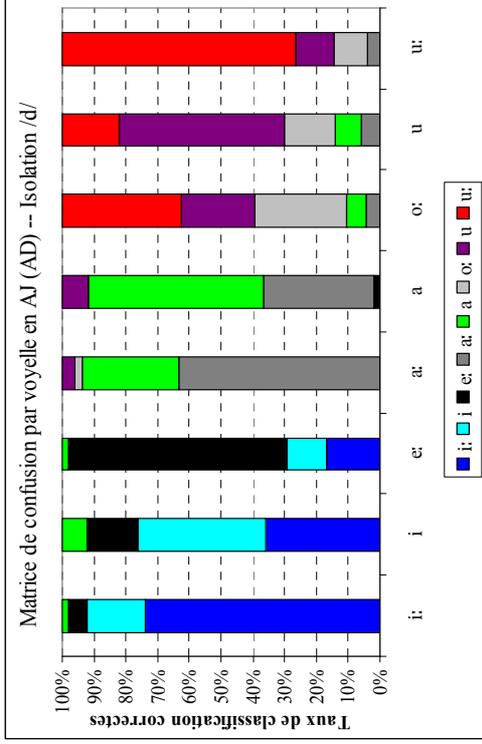
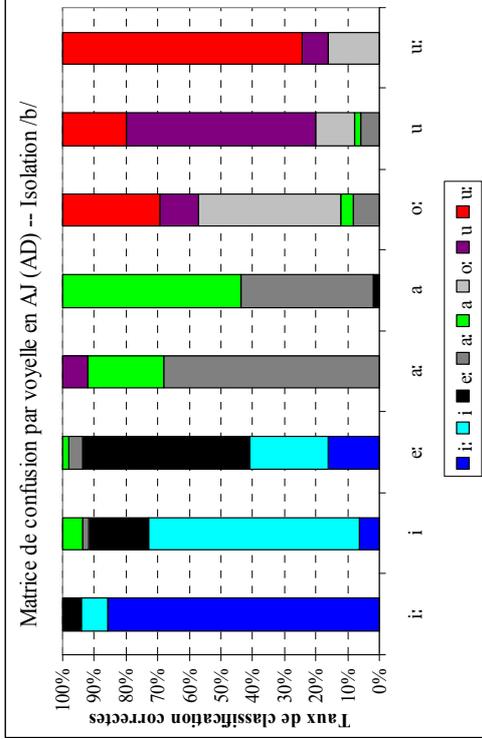




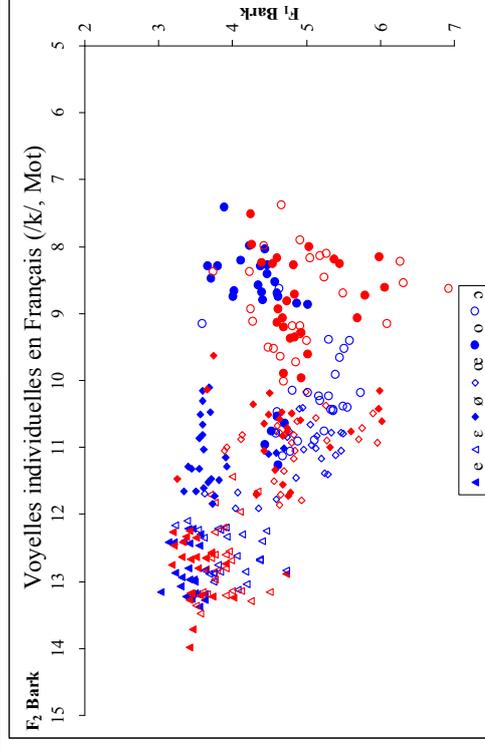
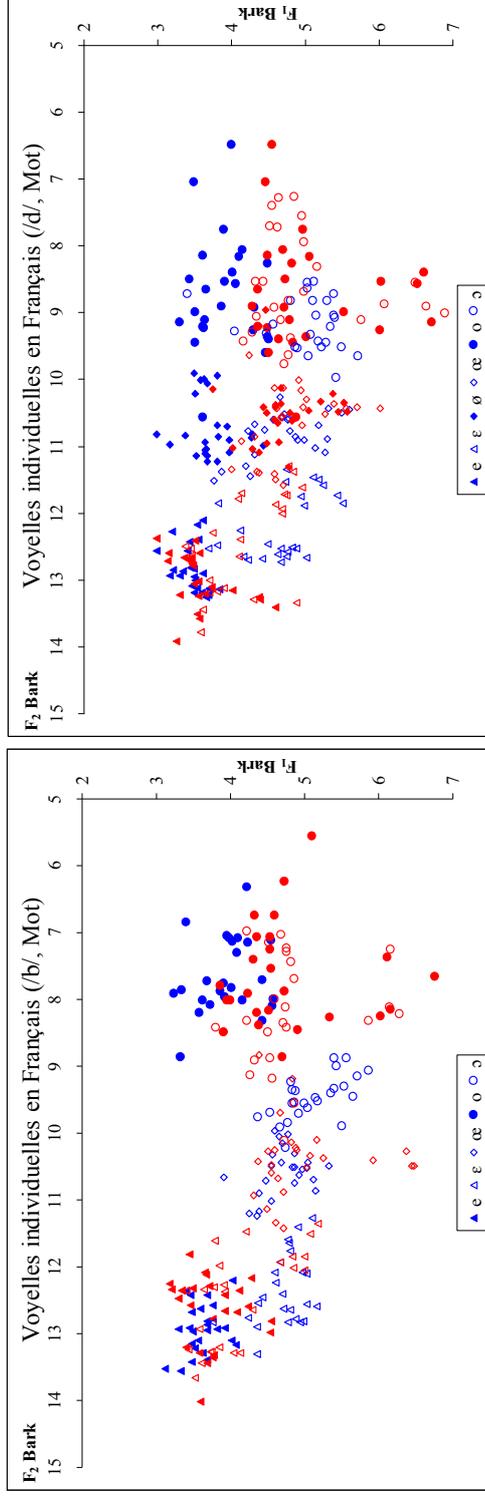
ANNEXE 25

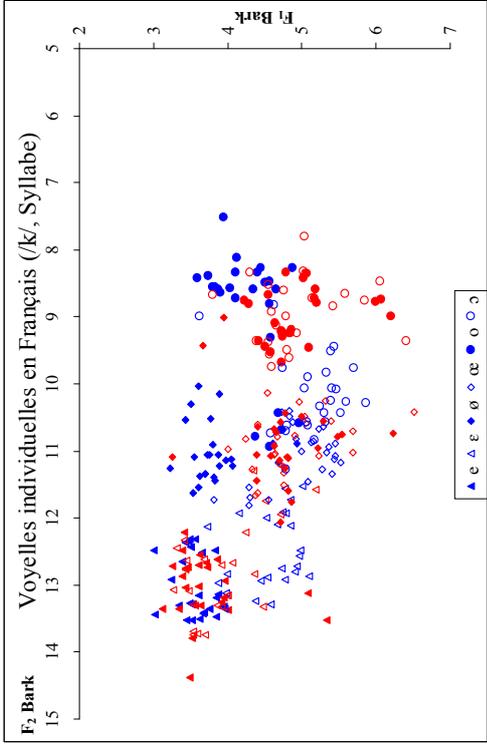
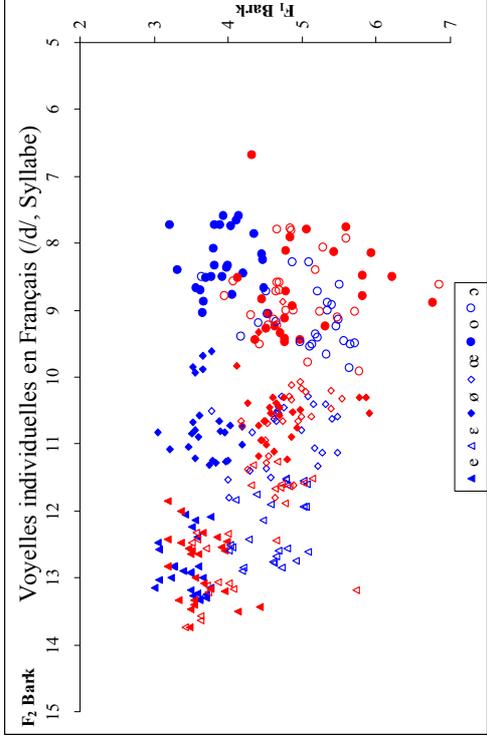
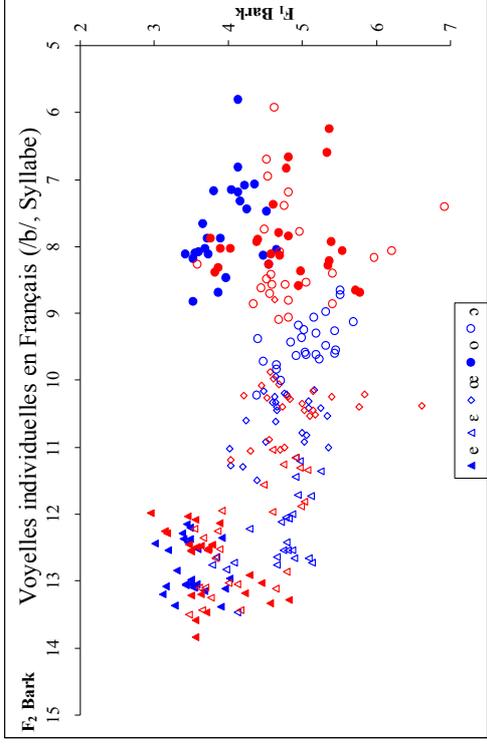
MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN ARABE JORDANIEN EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.

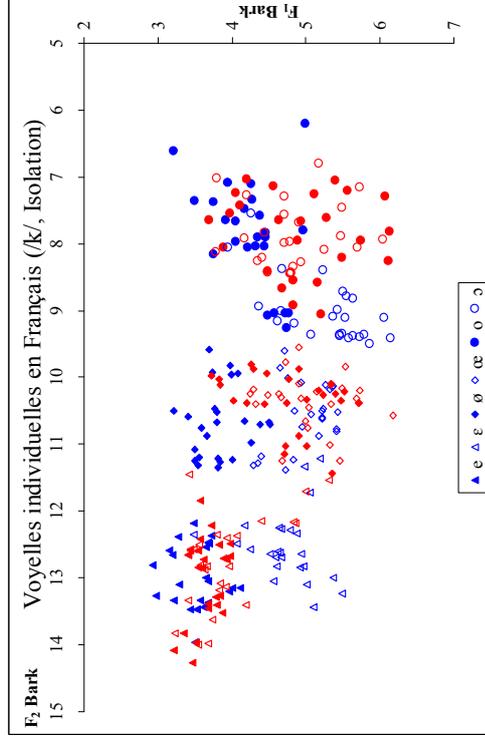
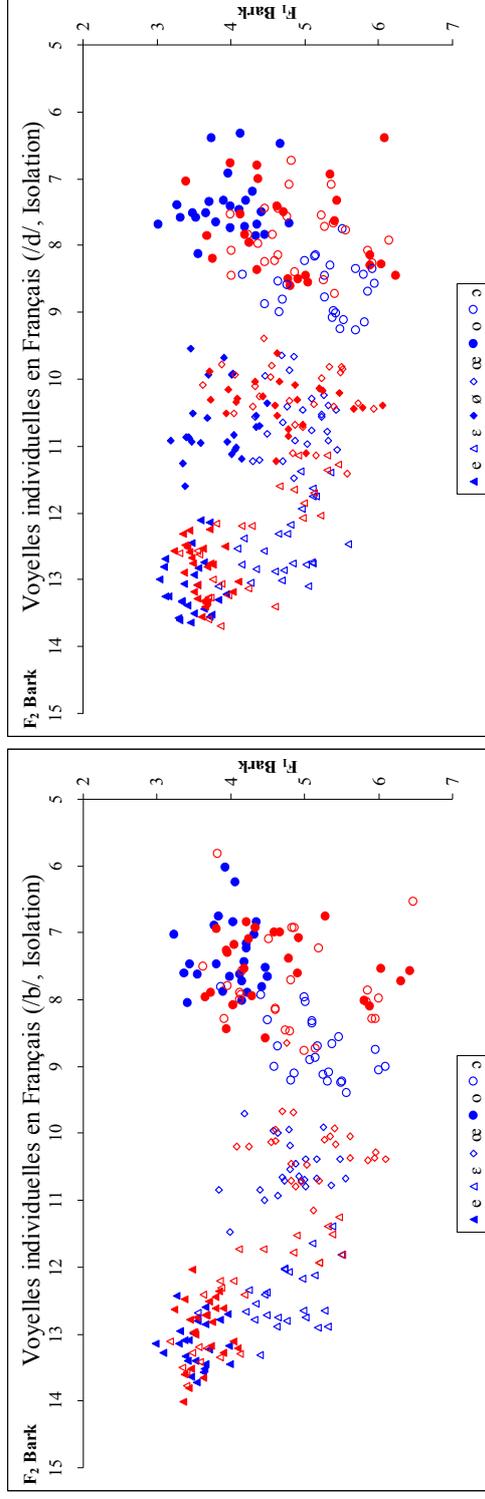




ANNEXE 26 VOYELLES INDIVIDUELLES EN FRANÇAIS SELON LES RÉGIONS : CENTRE (EN BLEU) ET SUD (EN ROUGE)

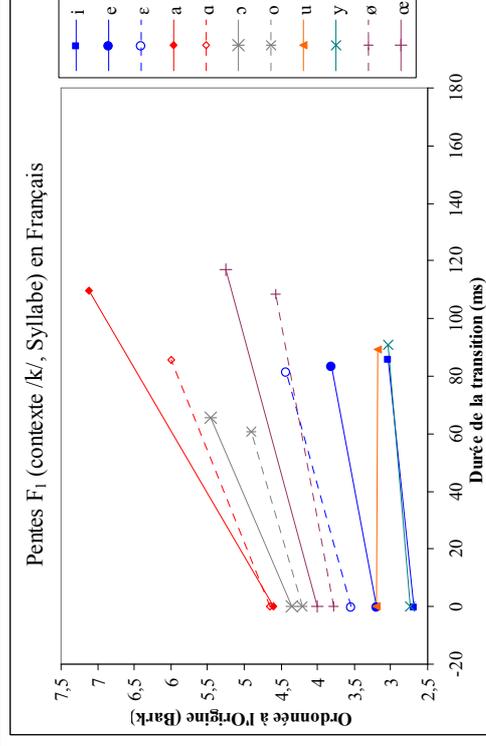
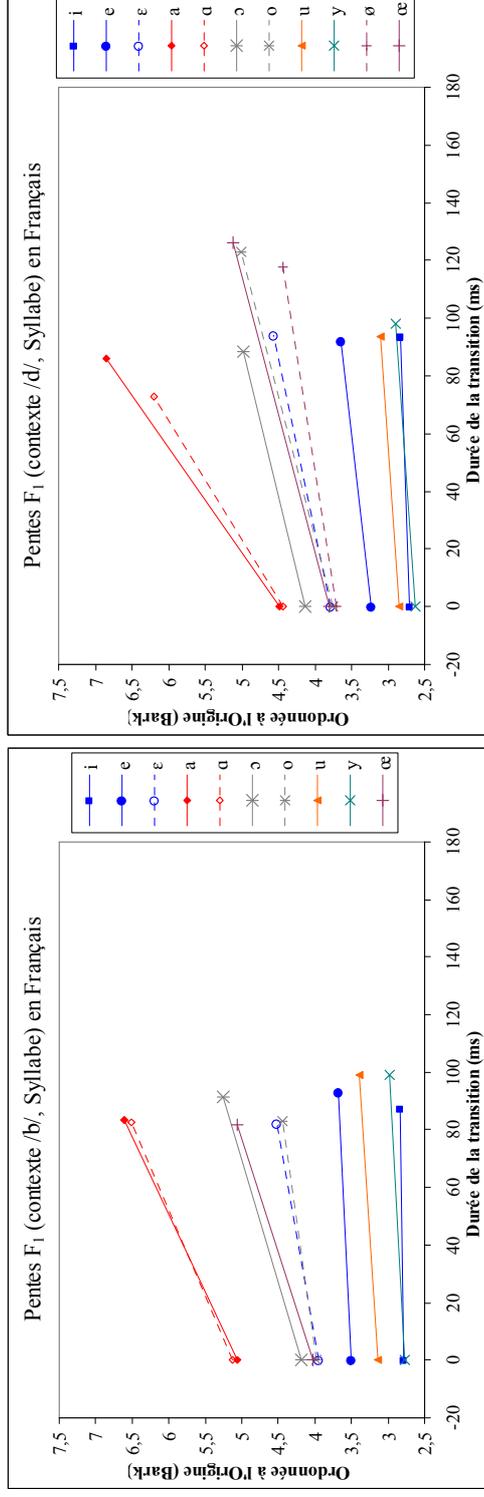


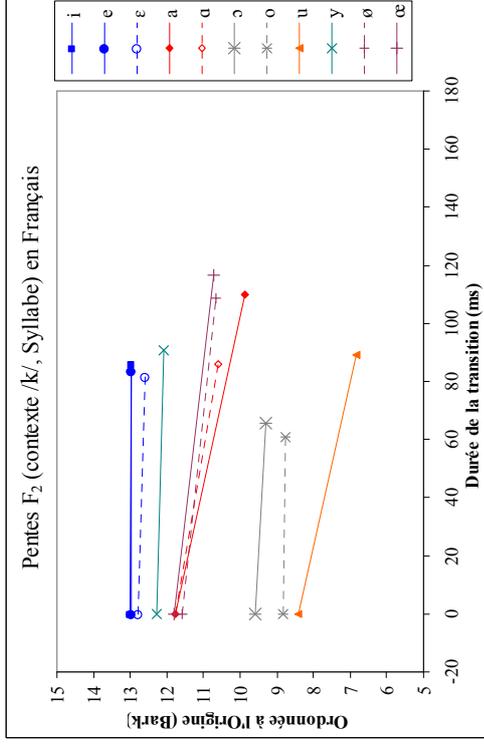
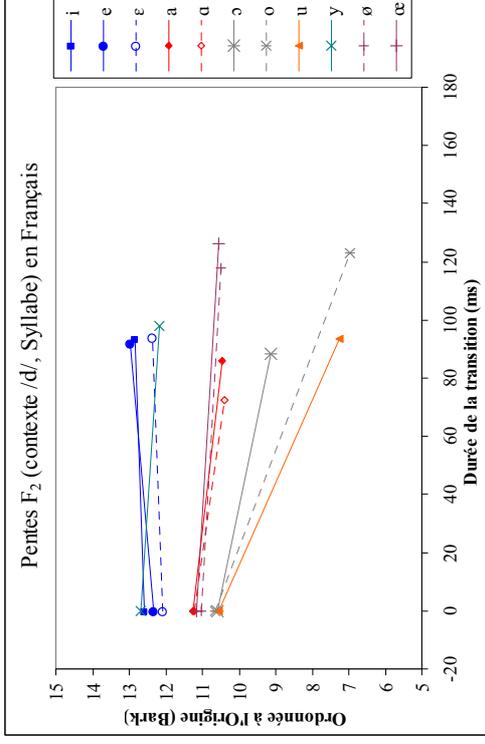
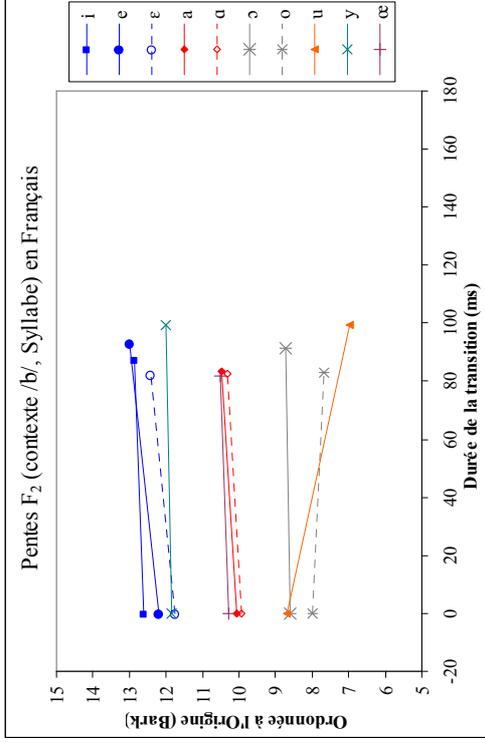


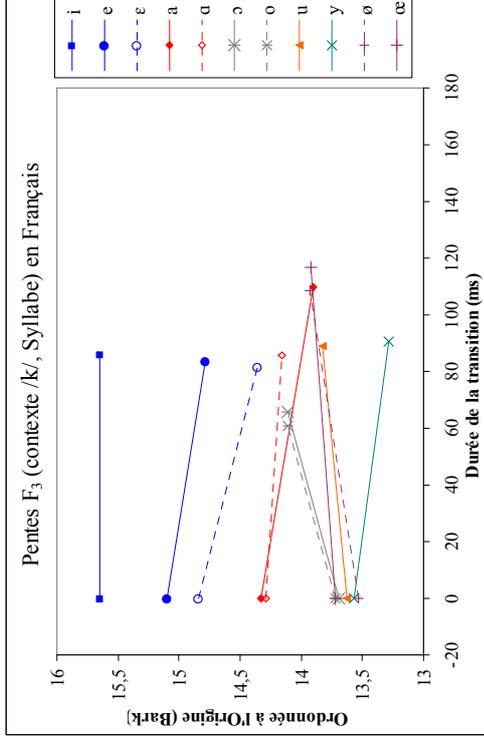
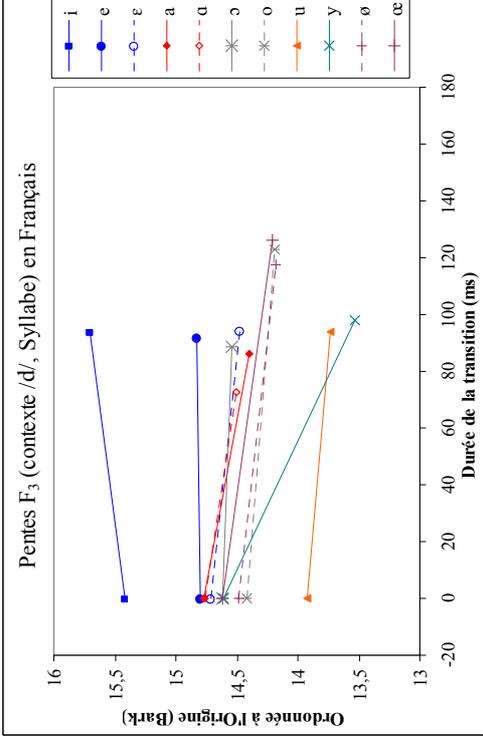
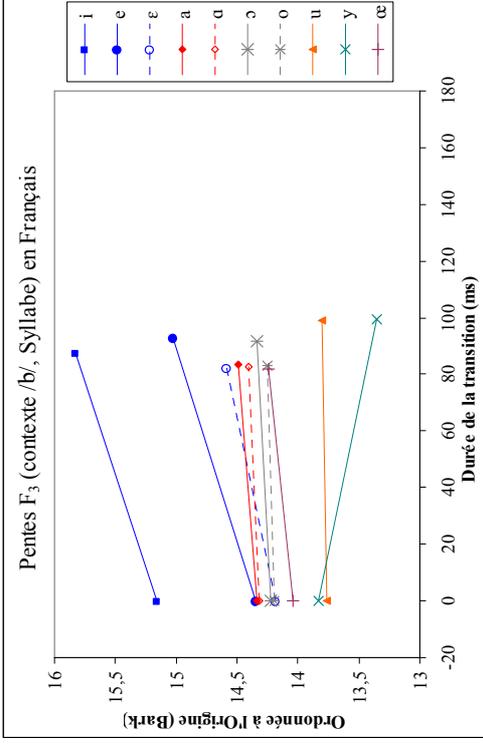


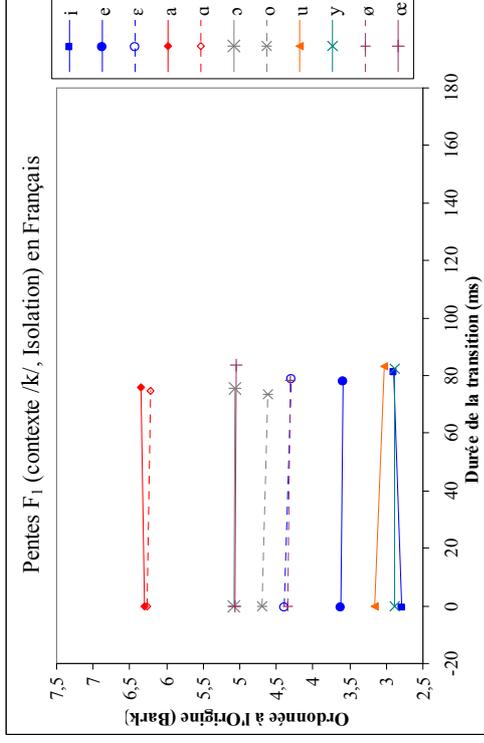
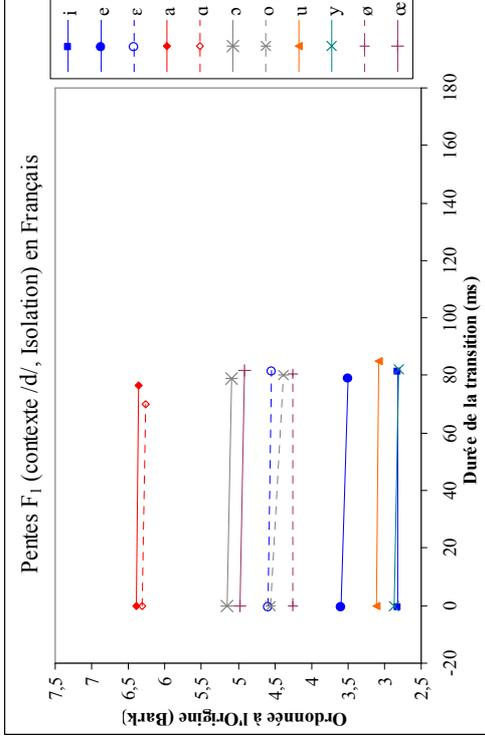
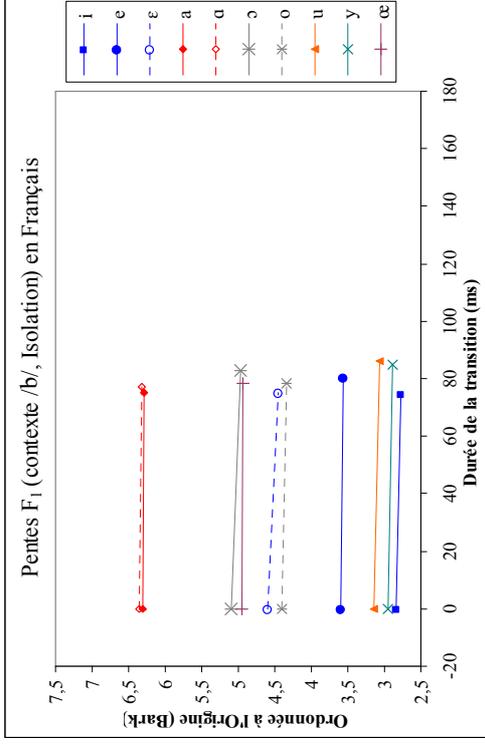
ANNEXE 27

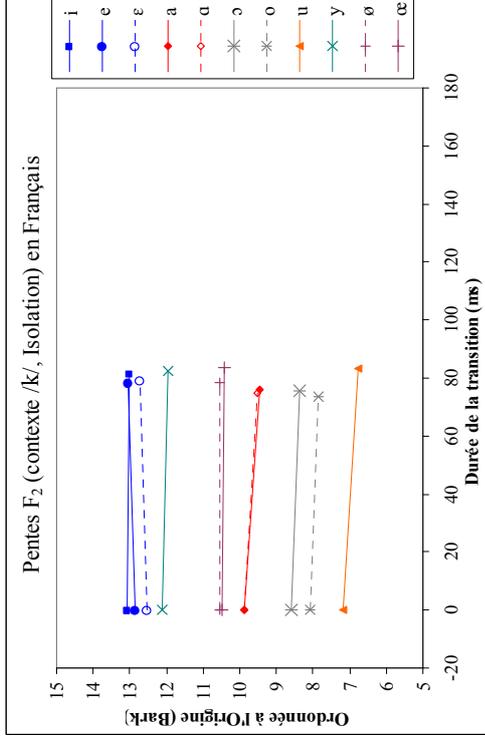
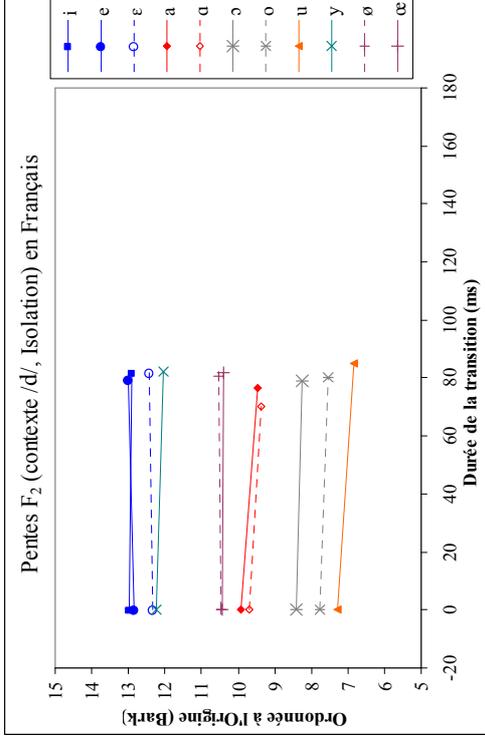
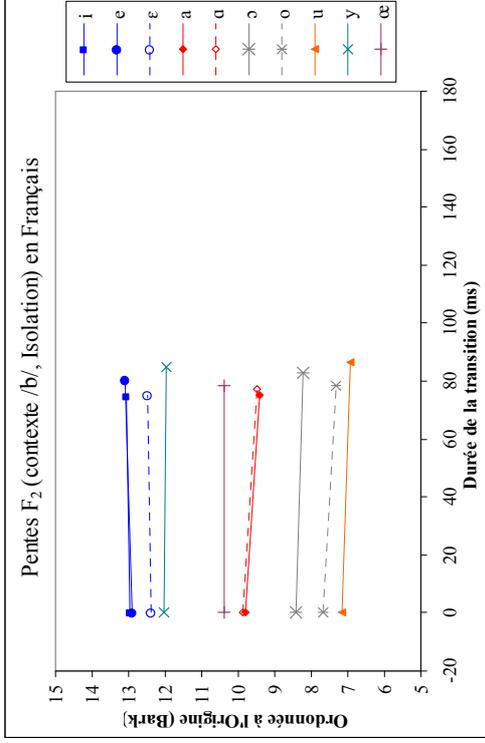
PENTES FORMANTIQUES DE F₁, F₂ & F₃ PAR VOYELLE EN SITUATIONS SYLLABE & ISOLATION EN FRANÇAIS.

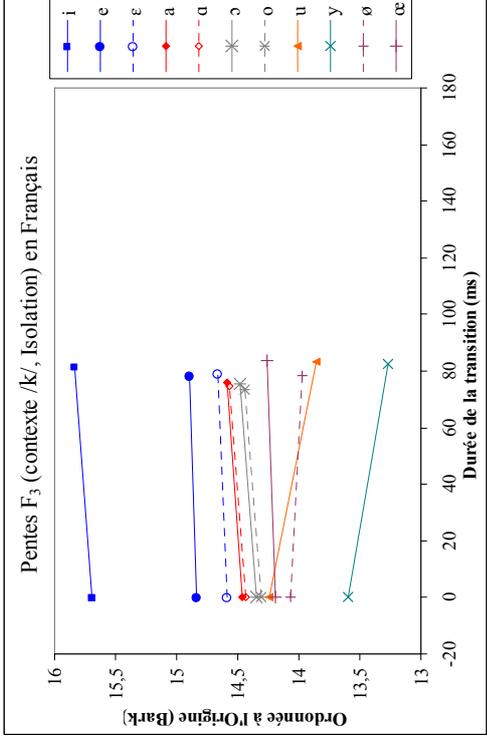
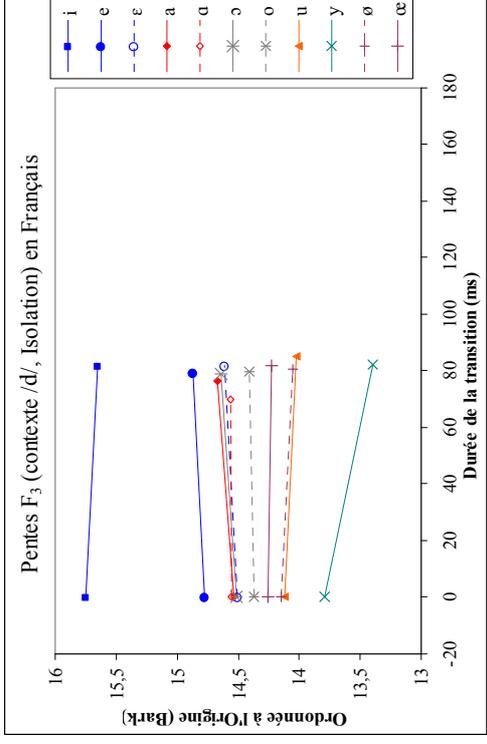
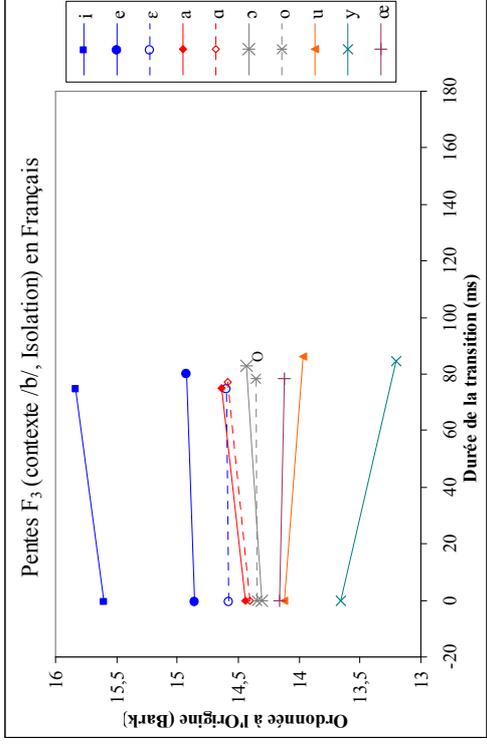








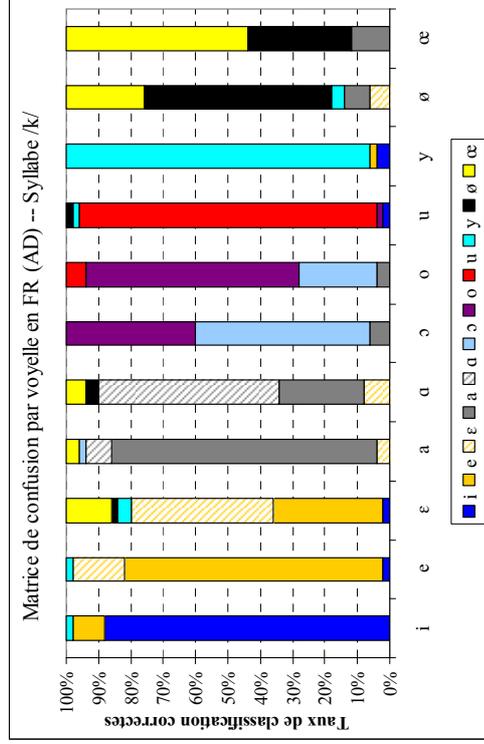
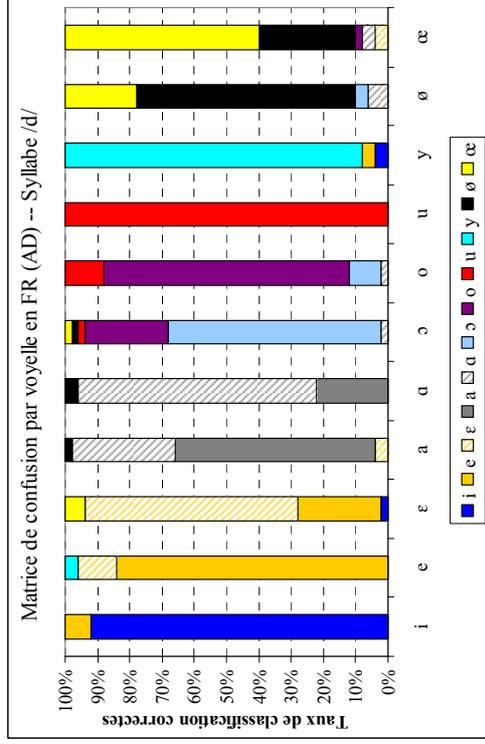
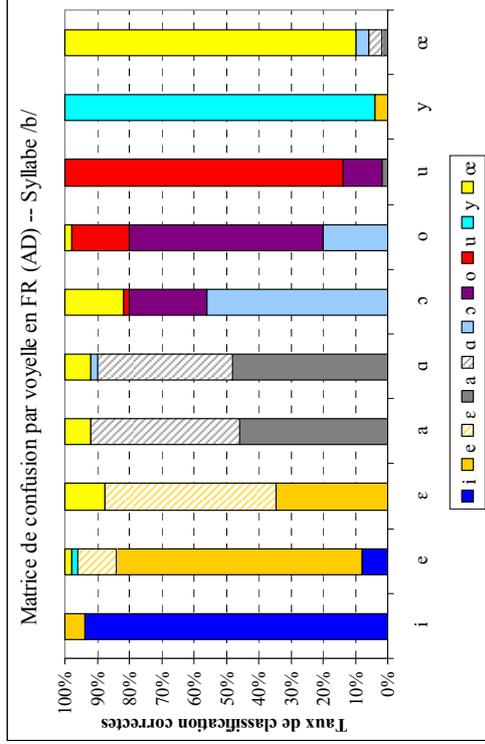


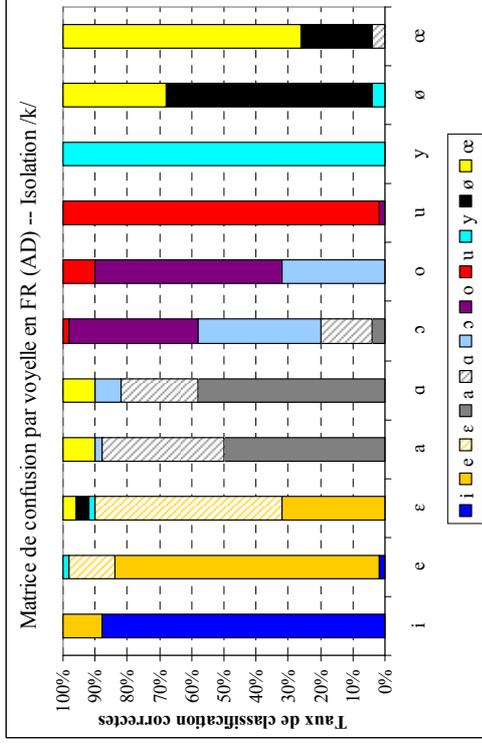
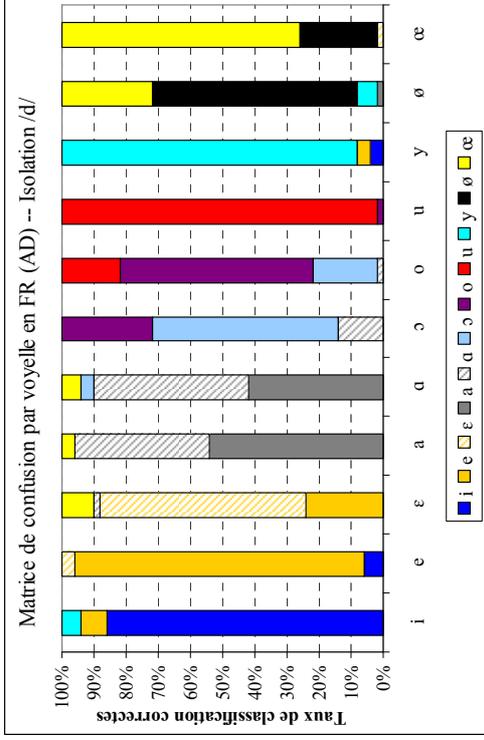
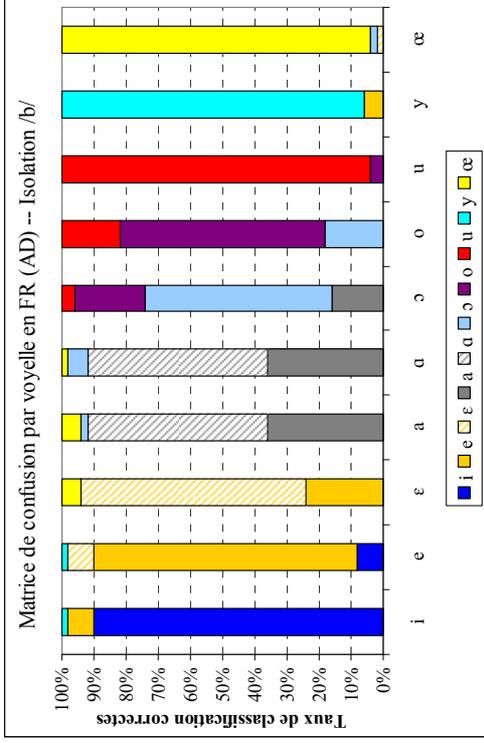


ANNEXE 28

MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN STATIQUE AVEC LA DURÉE

VOCALIQUE (AD) EN SITUATIONS SYLLABE ET ISOLATION.

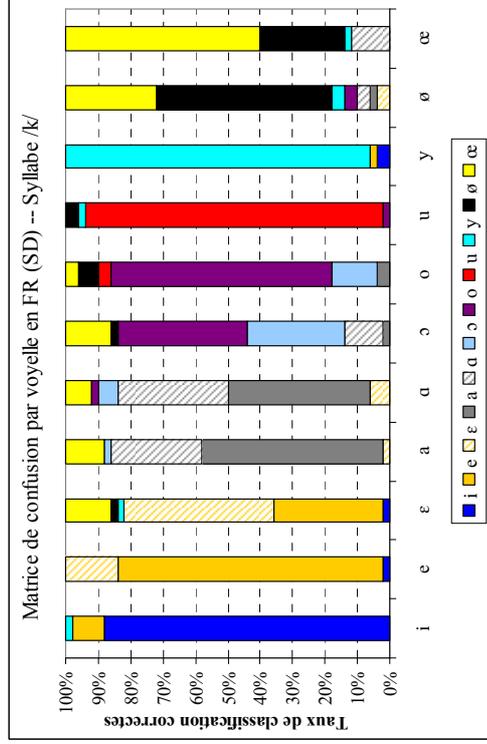
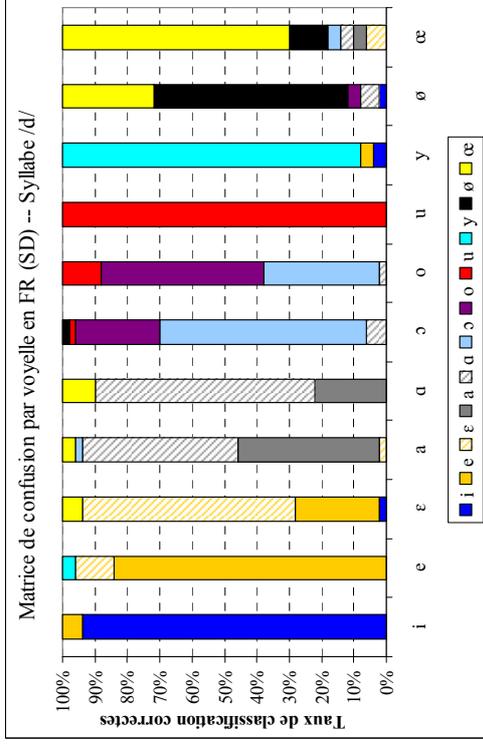
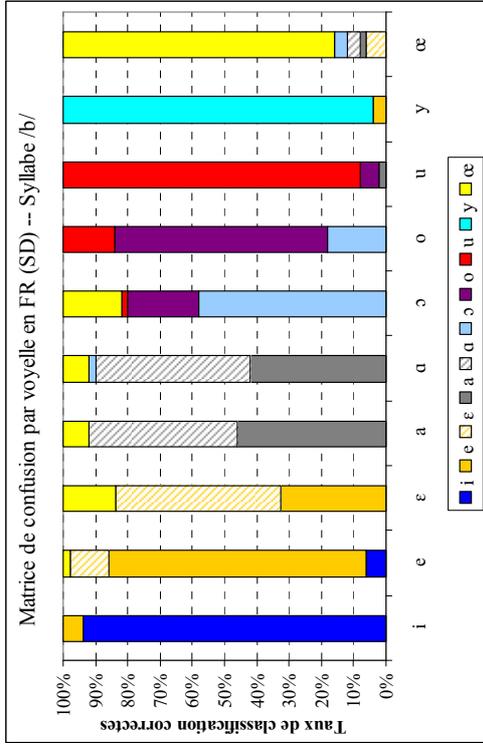


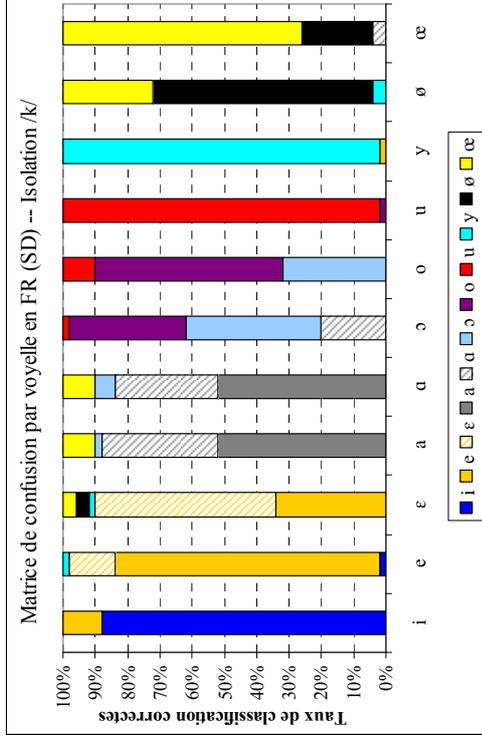
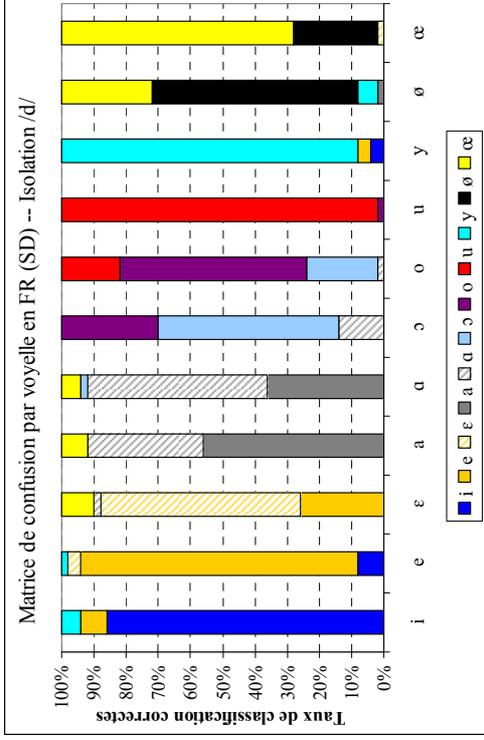
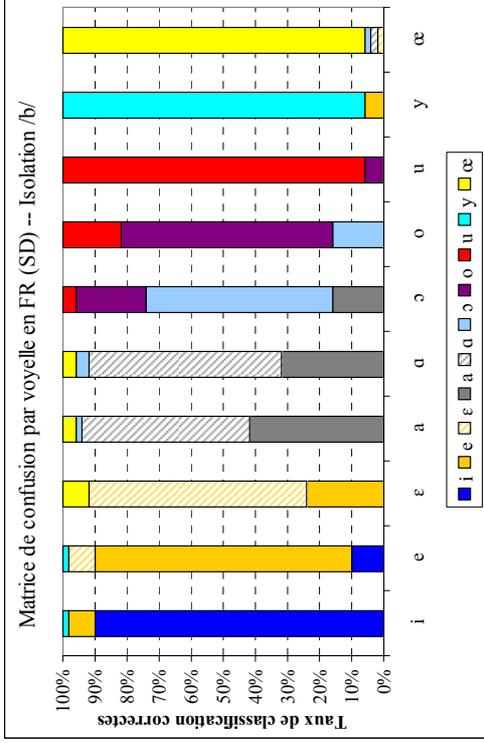


ANNEXE 29

MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN STATIQUE SANS LA DURÉE

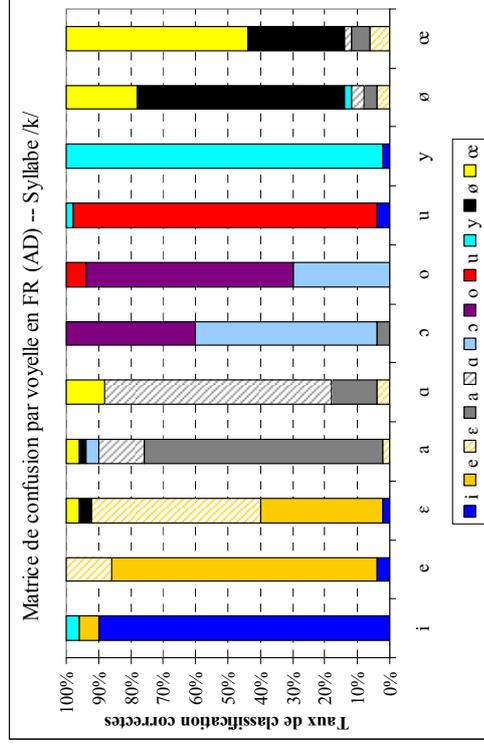
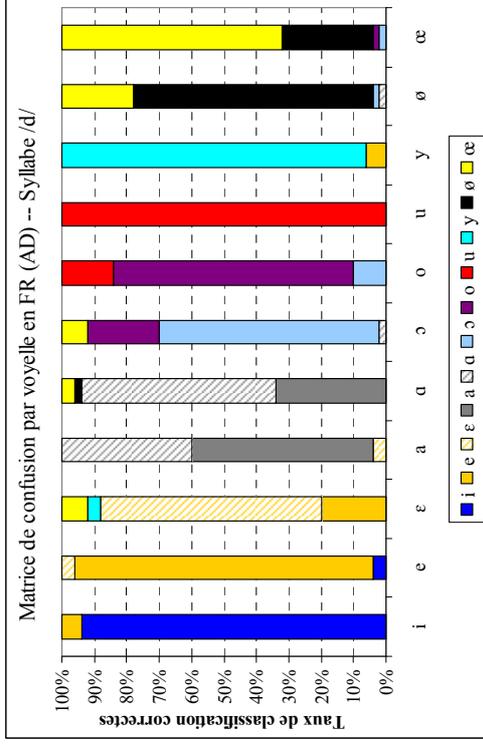
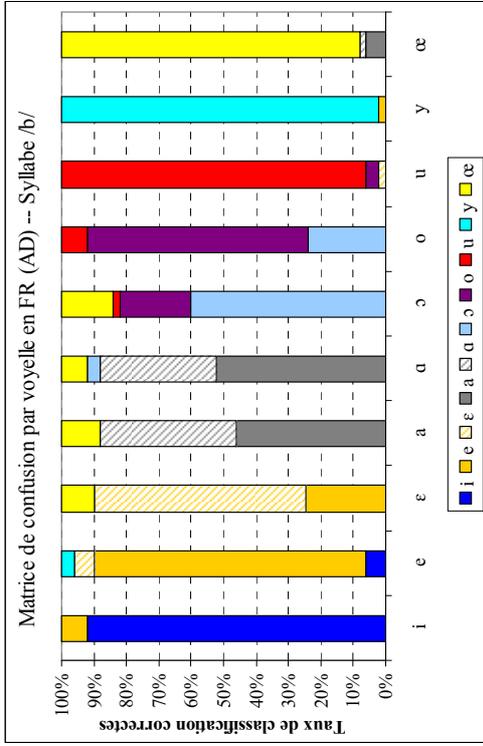
VOCALIQUE (SD) EN SITUATIONS SYLLABE ET ISOLATION.

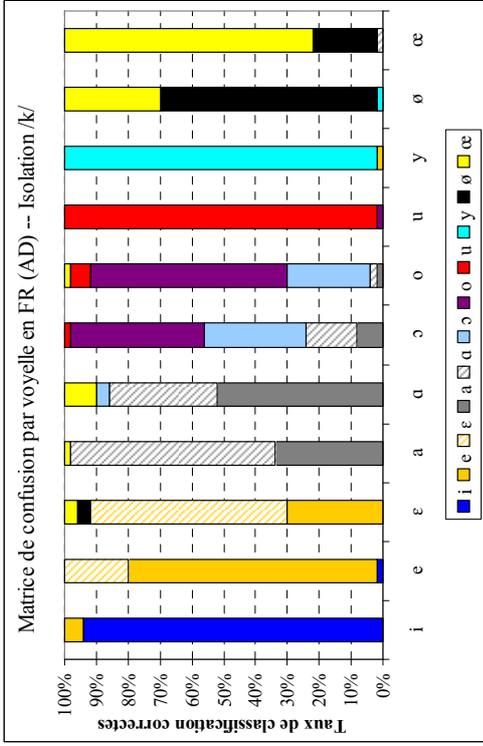
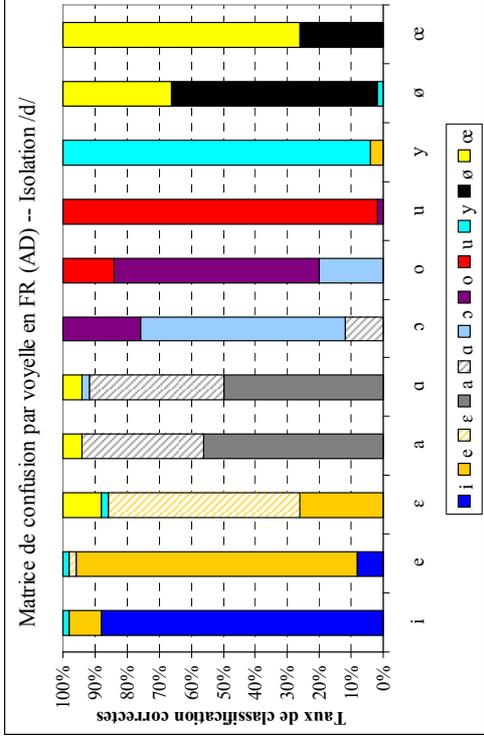
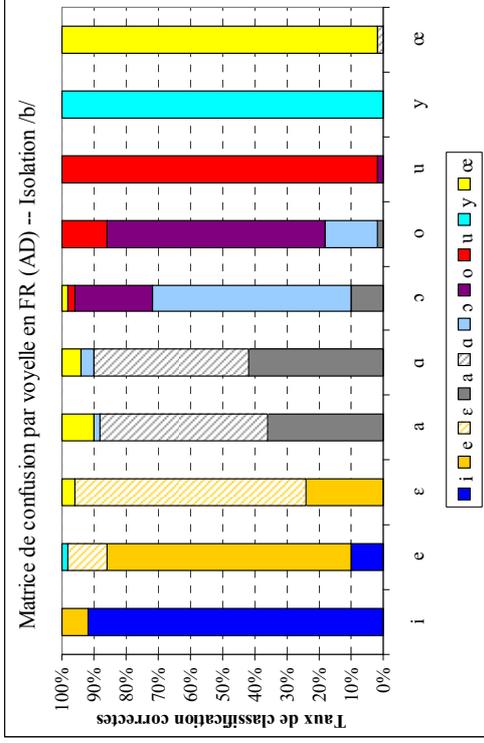




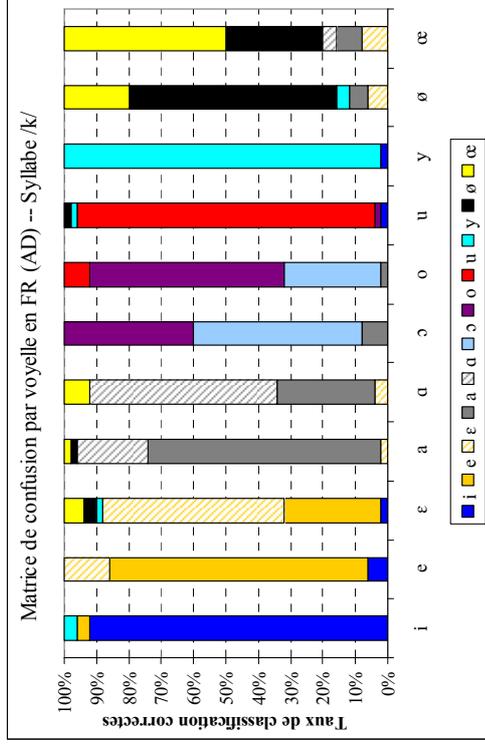
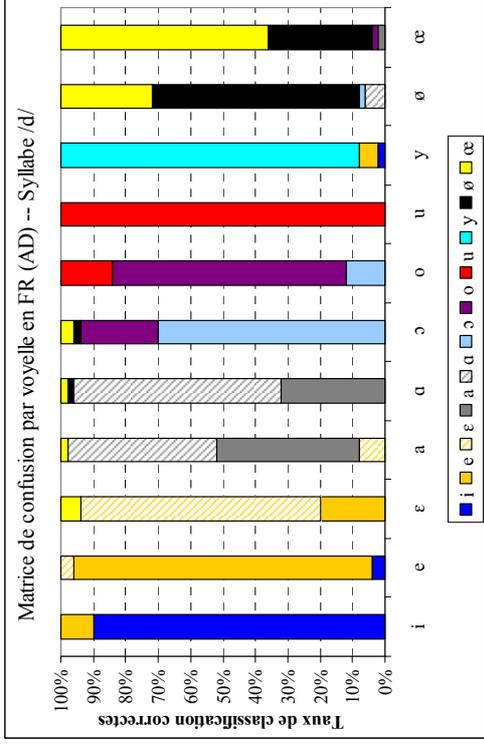
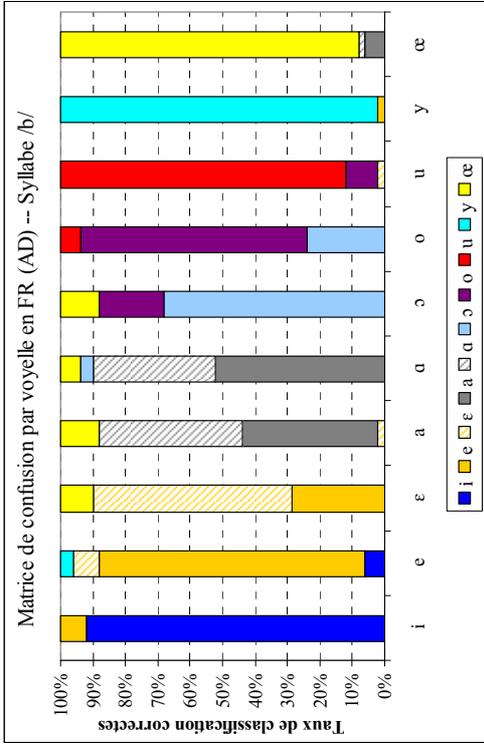
ANNEXE 30

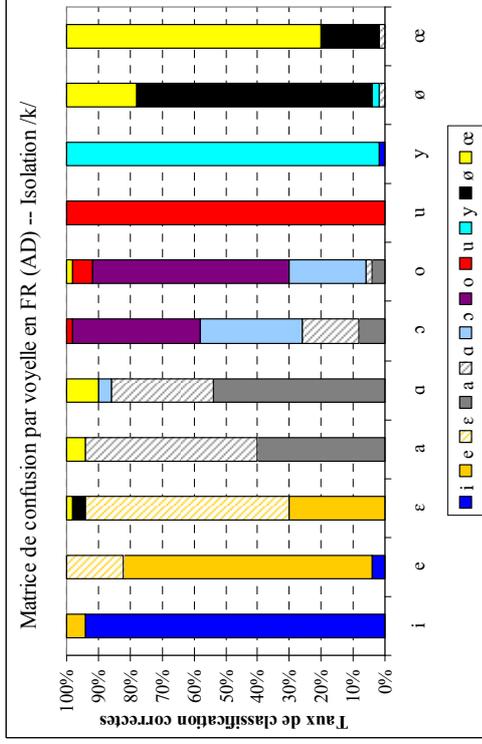
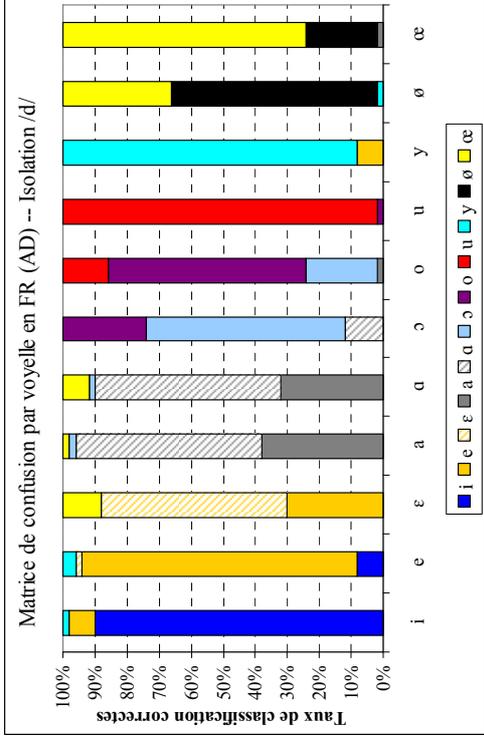
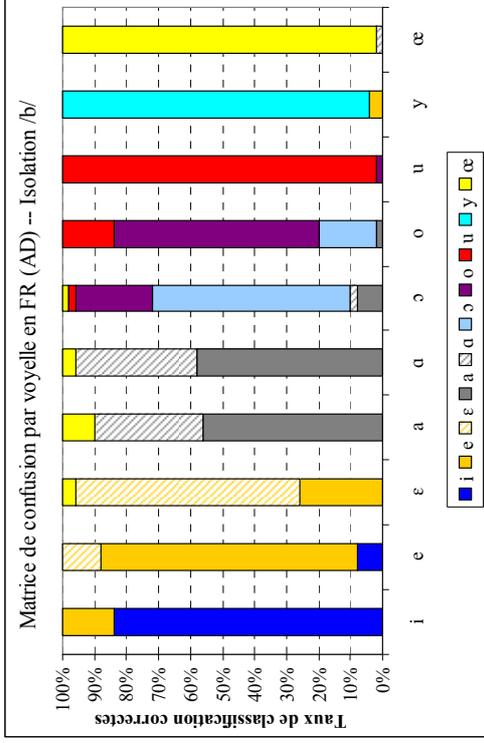
MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN DYNAMIQUE (RÉGRESSION LINÉAIRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.



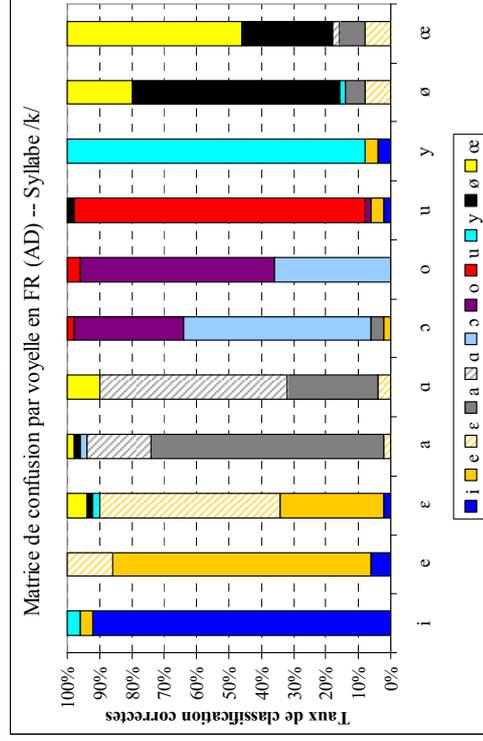
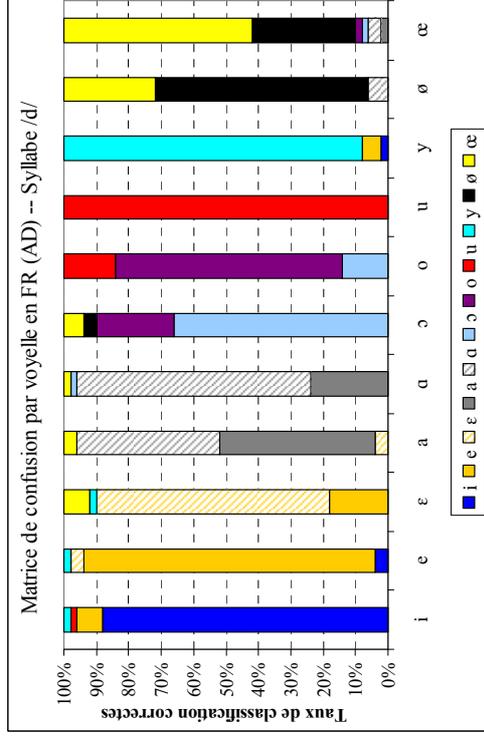
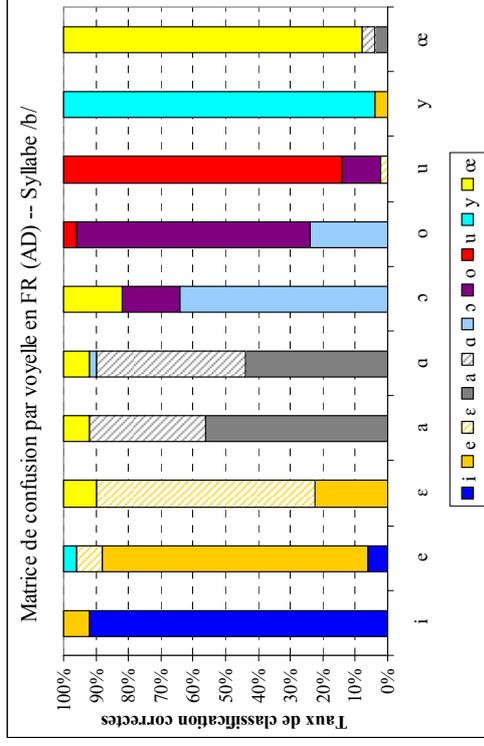


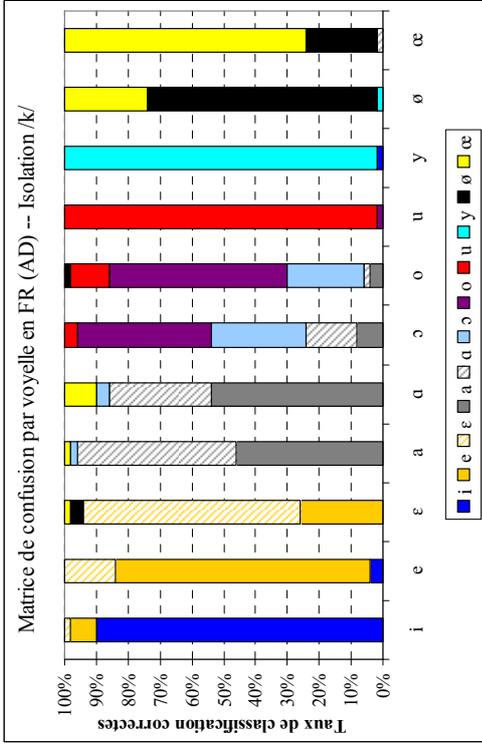
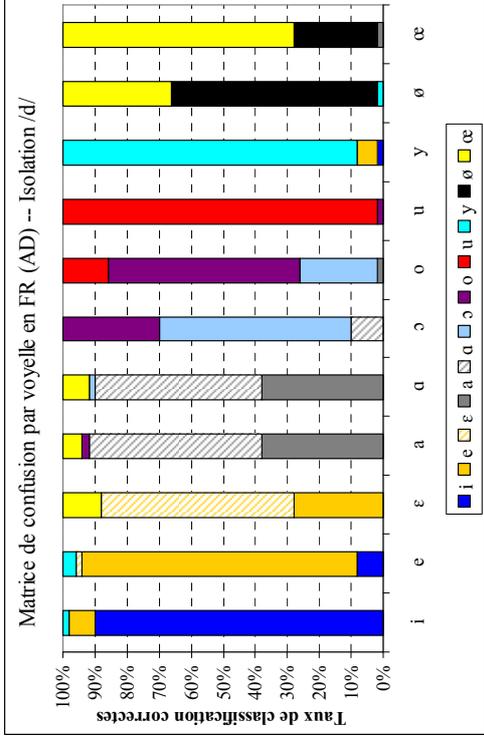
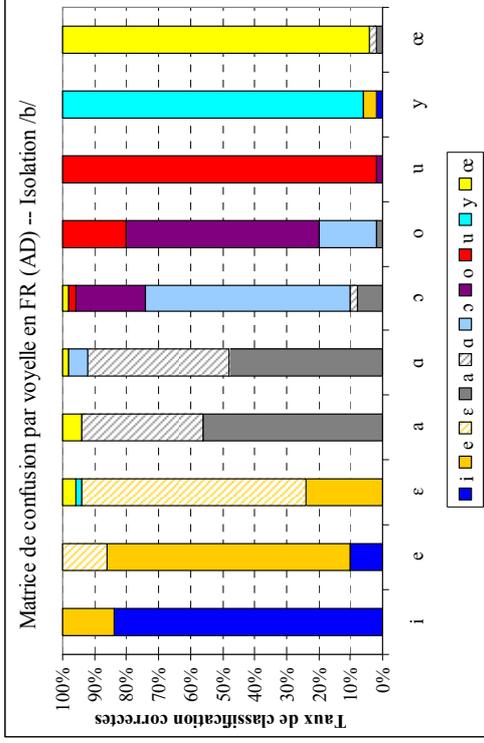
ANNEXE 31 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 2^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.





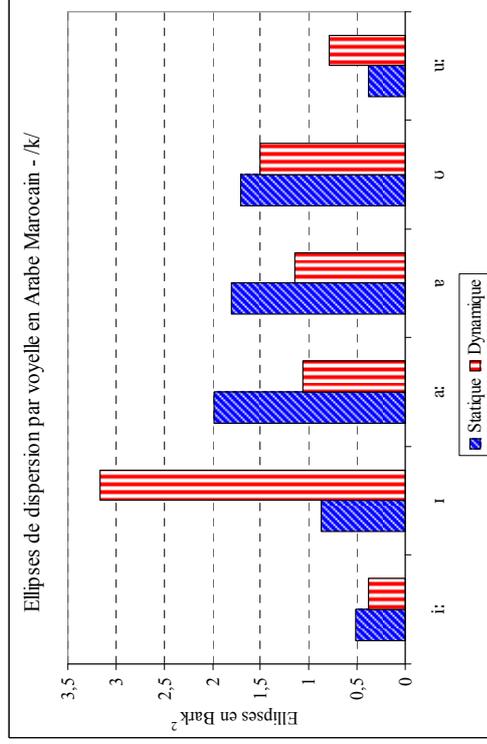
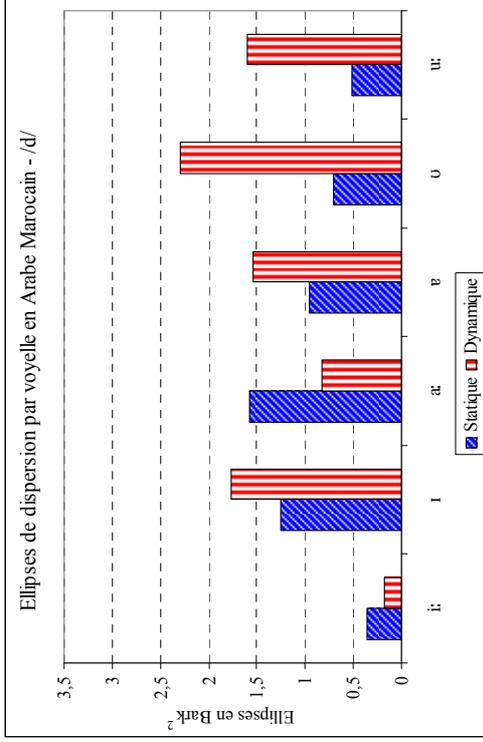
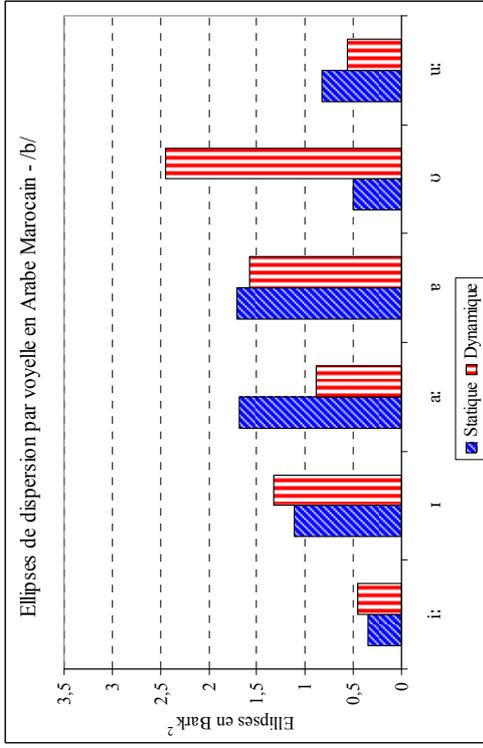
ANNEXE 32 MATRICES DE CONFUSION DE CLASSIFICATION DES VOYELLES EN FRANÇAIS EN DYNAMIQUE (RÉGRESSION POLYNOMIALE 3^{ÈME} ORDRE) AVEC LA DURÉE DE LA TRANSITION (AD) EN SITUATIONS SYLLABE ET ISOLATION.





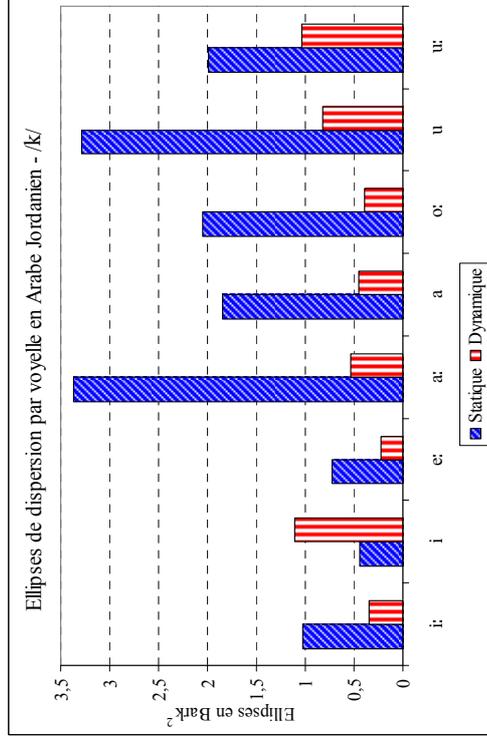
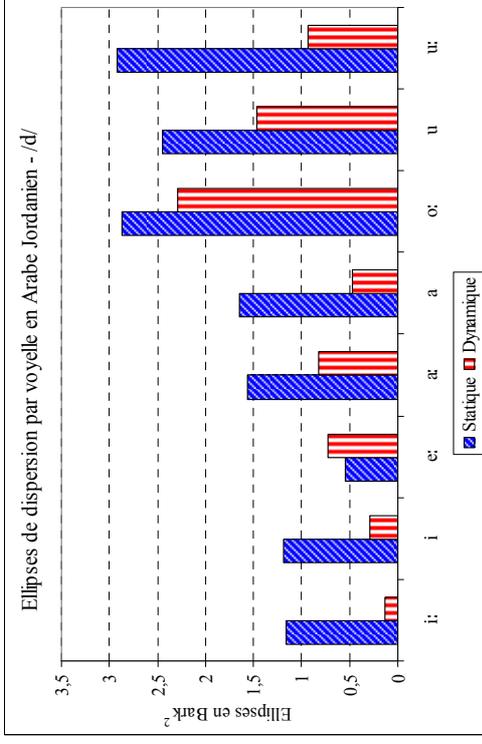
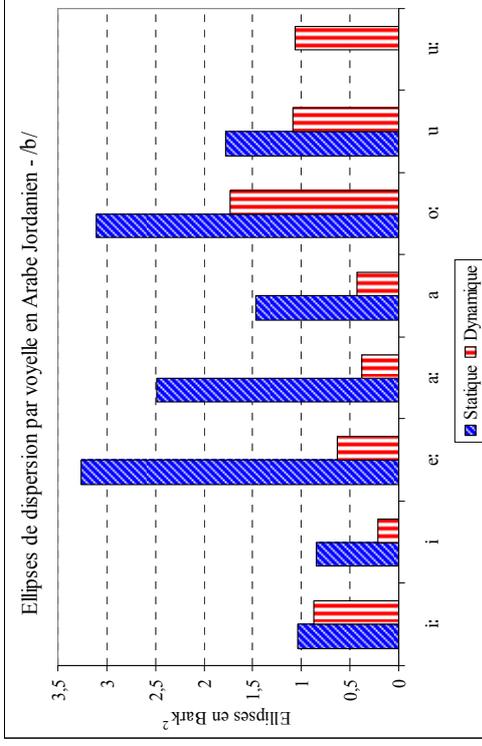
ANNEXE 33

ELLIPSES DE DISPERSION DES PROTOTYPES VOCALIQUES EN ARABE MAROCAIN EN FONCTION DU LIEU D'ARTICULATION



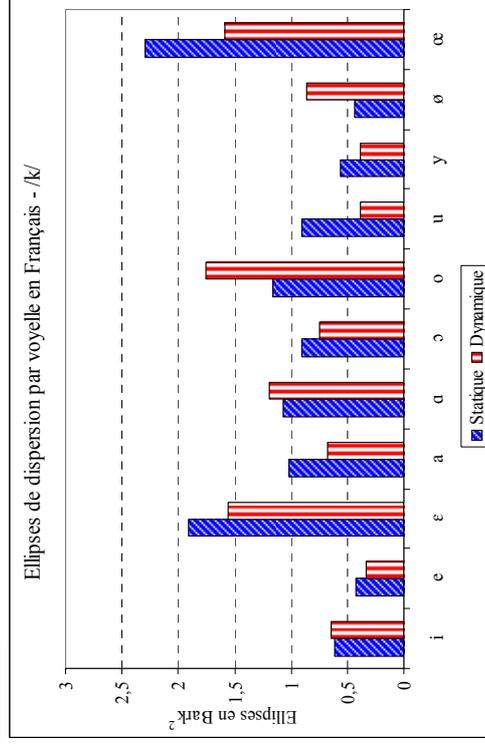
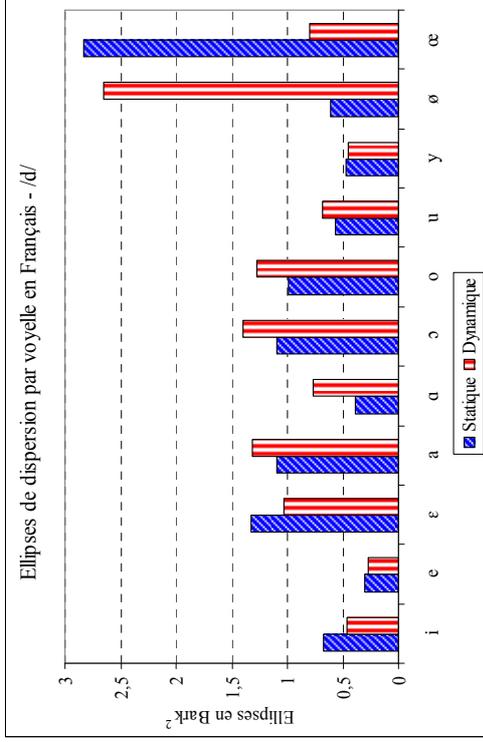
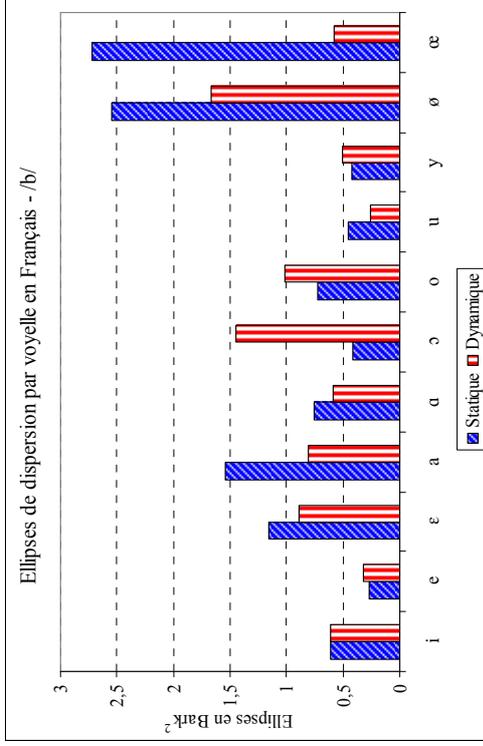
ANNEXE 34

ELLIPSES DE DISPERSION DES PROTOTYPES VOCALIQUES EN ARABE JORDANIEN EN FONCTION DU LIEU D'ARTICULATION



ANNEXE 35

ELLIPSES DE DISPERSION DES PROTOTYPES VOCALIQUES EN FRANÇAIS EN FONCTION DU LIEU D'ARTICULATION



« Indices dynamiques et perception des voyelles : étude translinguistique en arabe dialectal et en français »

Cette thèse a pour double objectif de comprendre l'impact des indices statiques et dynamiques en perception des voyelles en arabe et de comparer le rôle de la densité des systèmes sur la dispersion des voyelles en arabe marocain, arabe jordanien et français. Nous examinons le rôle de ces indices dans la discrimination entre les voyelles, en production et en perception (dans différents lieux d'articulation et situations de production). En production, les résultats montrent que l'utilisation des indices dynamiques permet une meilleure discrimination entre les voyelles des trois langues et une meilleure différenciation des voyelles dans chaque langue. Les différences inter-langues vont dans le sens de la théorie de la dispersion adaptée, où une centralisation de l'espace vocalique en arabe marocain par rapport aux deux autres langues est observée. En perception, les résultats indiquent une amélioration de l'identification des voyelles en arabe par rapport au français en utilisant les indices dynamiques. Il existe également des différences importantes liées à la densité des systèmes : les espaces vocaliques en arabe marocain sont, comme en production, plus centralisés comparés à ceux en arabe jordanien ou en français. Outre le fait que l'utilisation des indices dynamiques est particulièrement utile en arabe, les auditeurs français arrivent également à mieux discriminer les voyelles de leur système en dynamique. Ces résultats prouvent la complémentarité entre indices statique et dynamique. Suivant les propositions de l'approche *Dynamic Specification*, il semblerait que le rôle important des indices dynamiques en perception des voyelles soit indépendant de la langue étudiée.

Mots Clés : Arabe dialectal, français, production et perception de la parole, indices statiques et dynamiques, densité des systèmes vocaliques.

« Dynamic Cues and Vowel Perception: A Translinguistic study in Dialectal Arabic and French »

This dissertation has a double objective. On the one hand, we try to understand the impact static and dynamic cues have on the perception of Arabic vowels, and on the other hand, to evaluate the role of vowel density on the vocalic space organisation in Moroccan Arabic, Jordanian Arabic and French. First, we evaluate the role of different cues for vowel discrimination in production and perception (obtained in different consonantal environments and production situations). Results in production show that the use of dynamic cues helps to discriminate the vowels between and within the three languages. Inter-language differences may be related to the *Adapted Vowel Dispersion Theory*, where the vowel space in Moroccan Arabic seems more centralised than that of Jordanian Arabic or French. At the perceptual level, results indicate that vowel identification in Arabic improves when dynamic cues are used in comparison with French. Some important differences between the vowel spaces in the three languages are shown. These may be caused by vowel density effects. Although the use of dynamic cues is more adapted to Arabic, French listeners were helped when dynamic cues were available. These results indicate that static and dynamic cues are complementary in vowel description between the languages. As suggested by the *Dynamic Specification* model, it seems that the role of dynamic cues in vowel perception is independent from the language under scrutiny.

Key Words: Dialectal Arabic, French, production and perception of speech, static and dynamic cues, vowel systems density.

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité - pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer ni l'adapter.