

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

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

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 α/ 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	fondue	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.

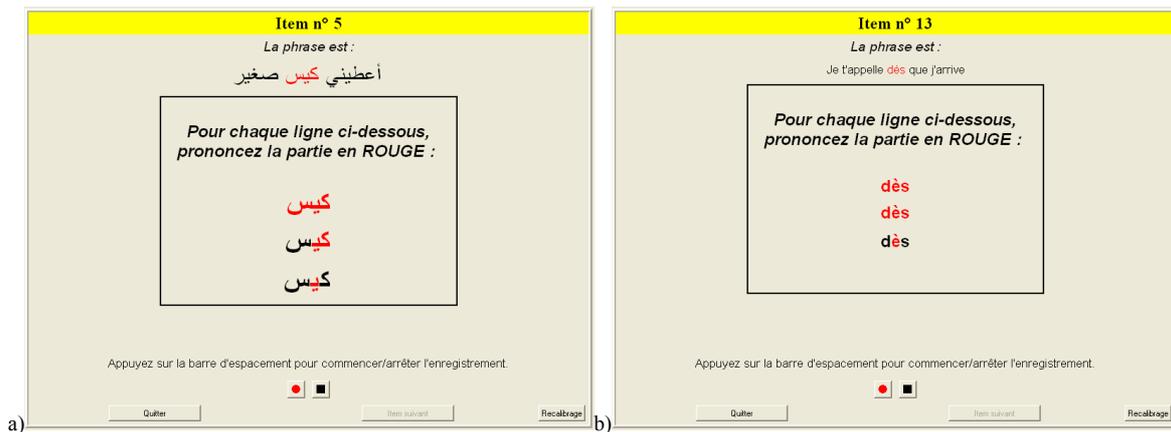


Figure 27 : Interfaces du logiciel dans la tâche d'enregistrement des voyelles en AJ (partie a) et en FR (partie b), présentant la phrase porteuse avec l'item à produire. Les parties que les locuteurs doivent produire en fonction des situations de production apparaissent en rouge.

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.

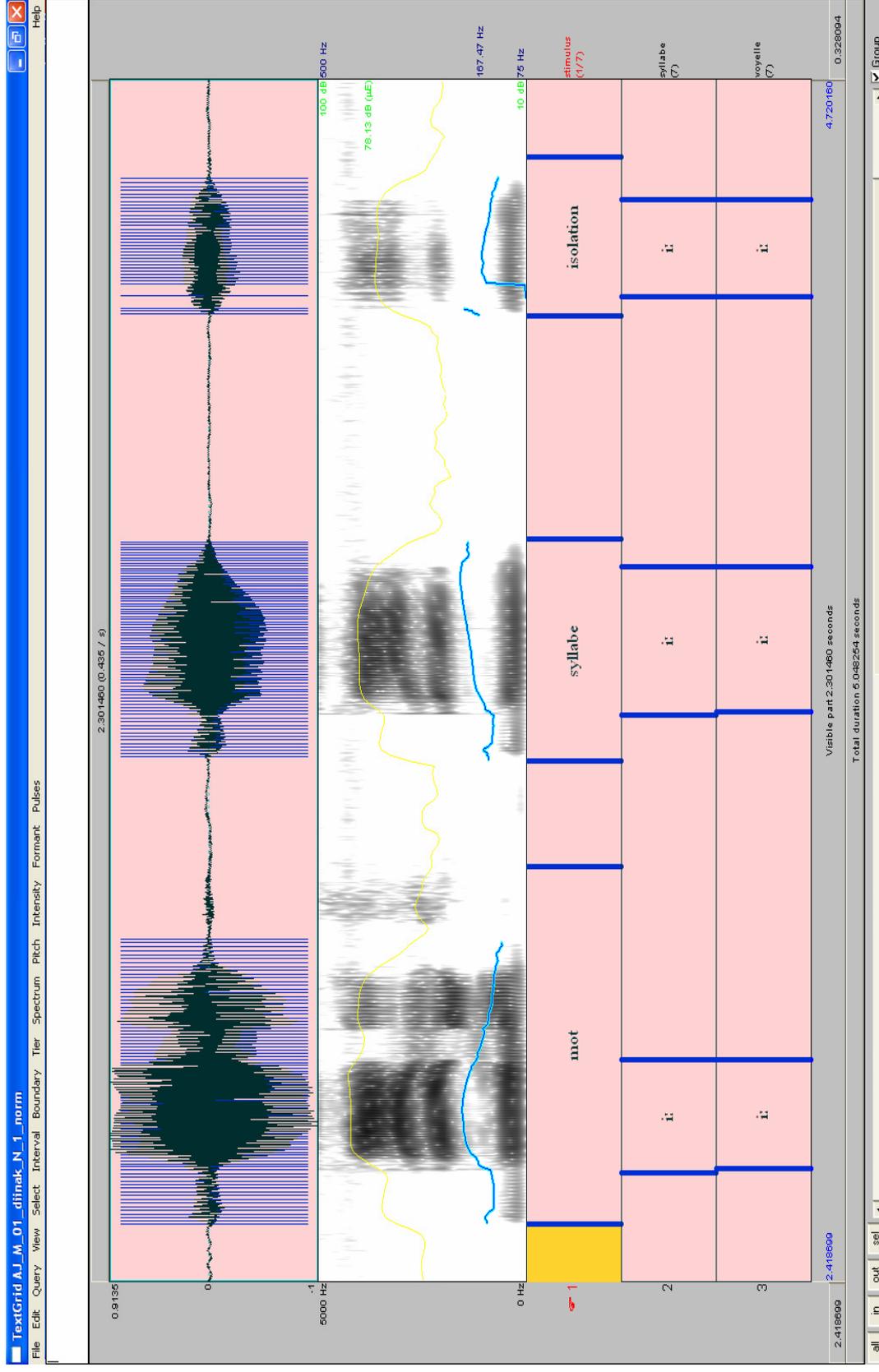


Figure 28 : Segmentation du mot [di:nak] « ta religion » en AJ, produit par le locuteur AJM01 dans les trois situations de production.

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.

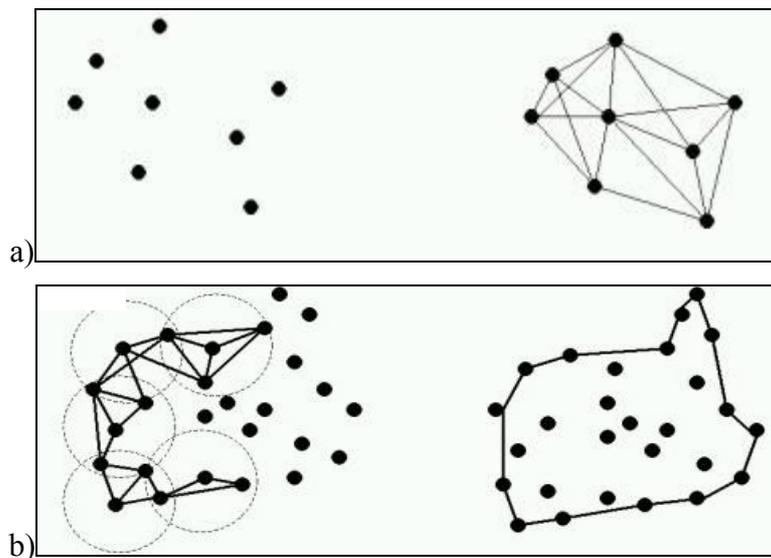


Figure 29 : Schémas montrant le principe de base du calcul de l'enveloppe convexe. En partant de plusieurs points (partie a), on calcule les différents triangles possibles, on les fusionne et on obtient un polygone (partie b).

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.

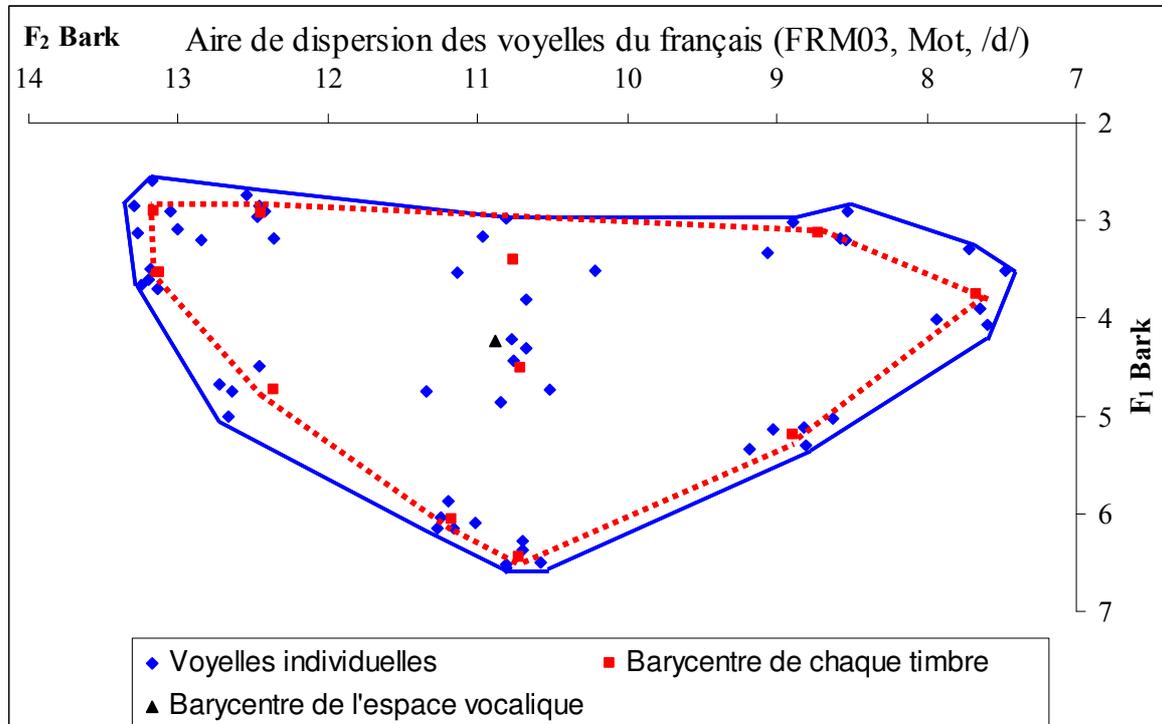


Figure 30 : Schéma représentant la dispersion des voyelles du FR produites par le locuteur FRM03 dans la situation de production Mot en contexte/d/ : en ligne bleue continue est représentée l'espace vocalique quadrilatère et en ligne rouge discontinus, l'Enveloppe Convexe de cet espace vocalique.

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.

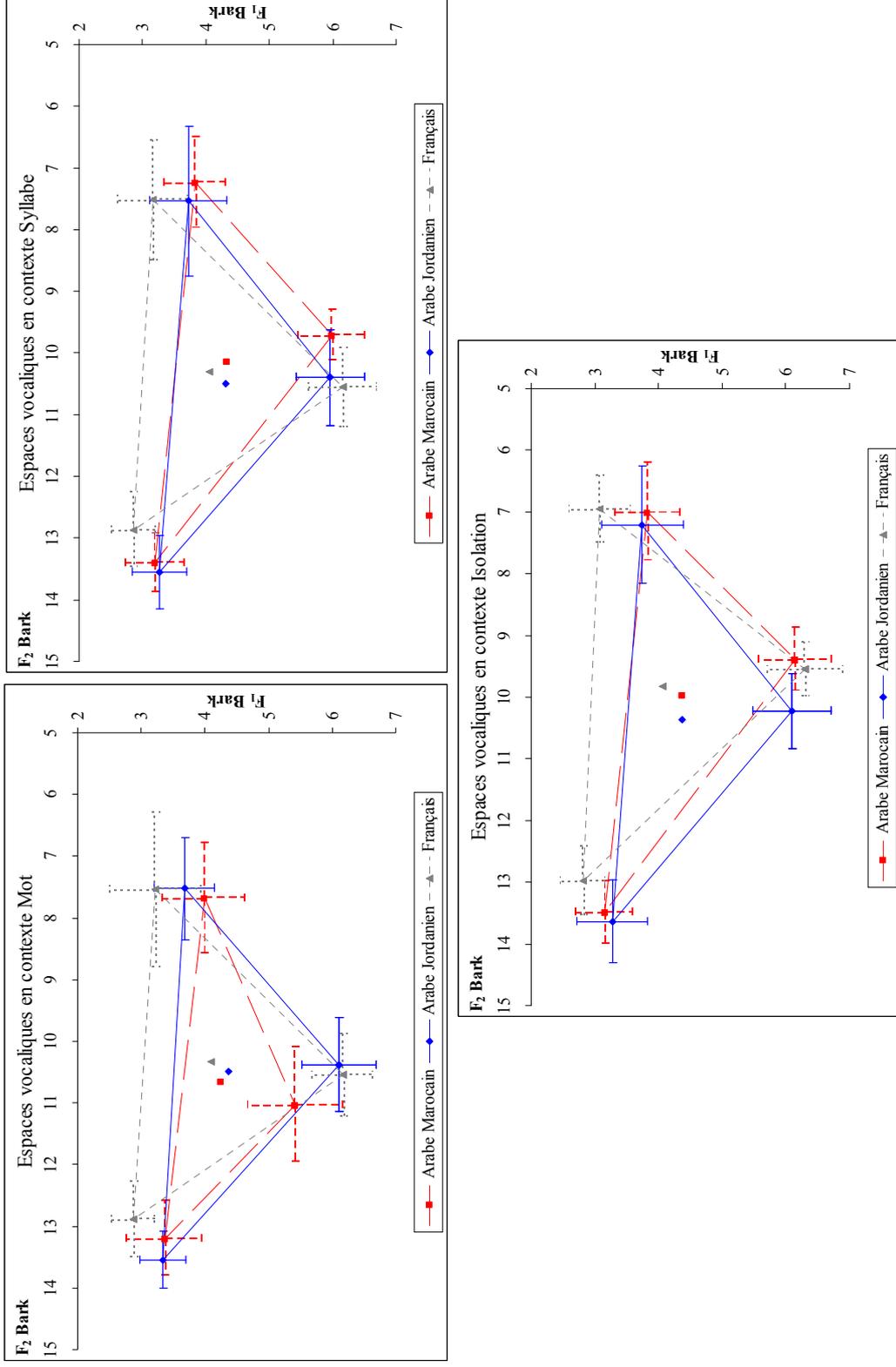


Figure 31 : Dispersion des trois timbres /i a u/ en AM, AJ et FR en fonction de la situation de production (voyelles produites dans les environnements /b d k/). Les points au centre des espaces vocaux représentent les Barycentres pour chaque langue.

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 corrélés 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.

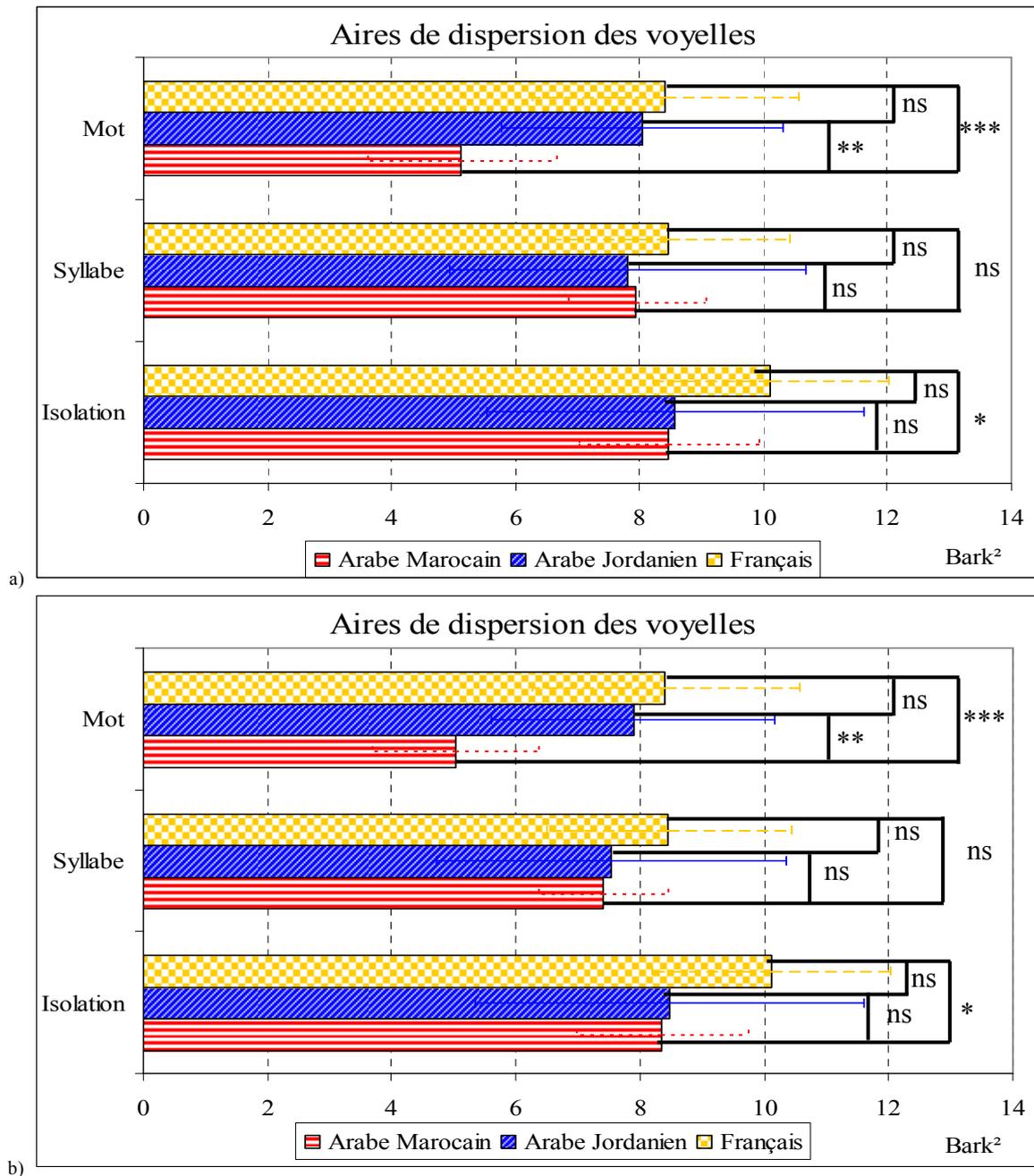


Figure 32 : Aires de dispersion (moyennes et écart-types) obtenues pour les voyelles /i a u/ en AM, AJ et FR en fonction de la situation de production, dans les environnements /b d k/ (partie a) et /b d d̥ k/ (partie b). (ns = non significatif ; * = p<0,05 ; ** = p<0,01 ; *** = p<0,001).

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.

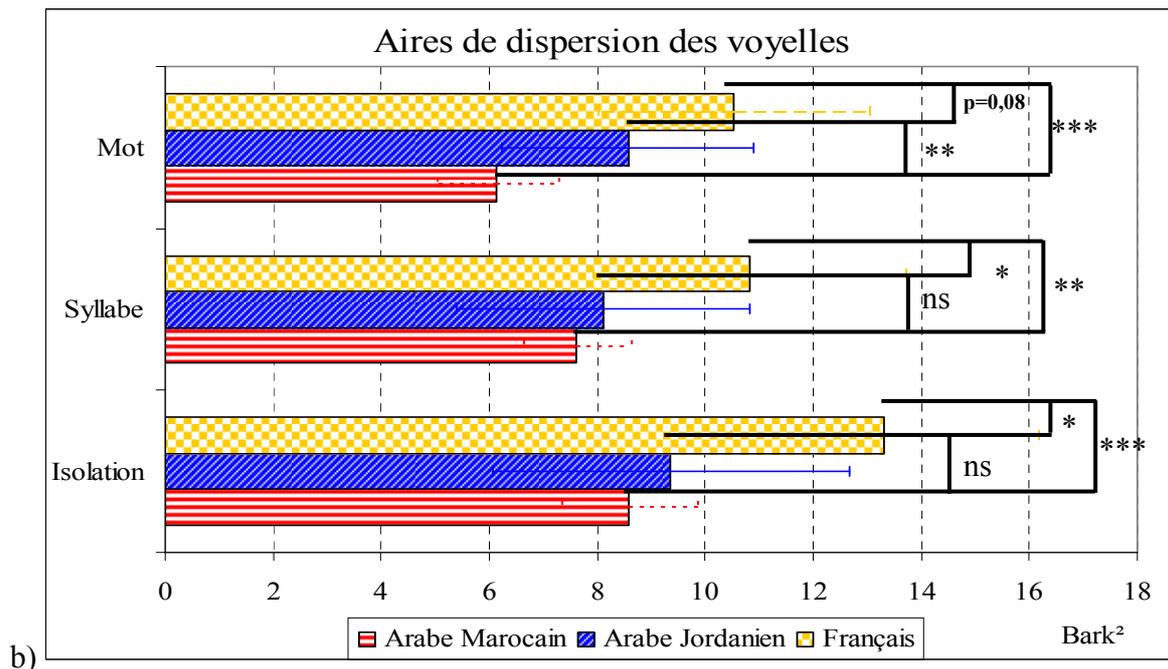
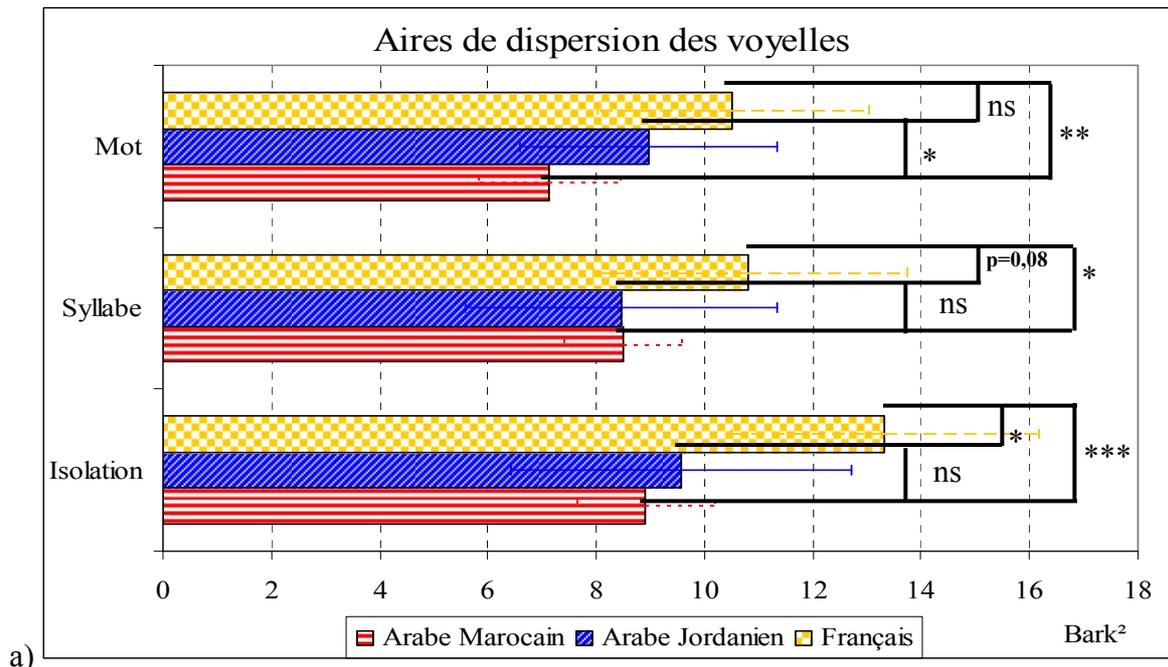


Figure 33 : Aires de dispersion (moyennes et écart-types) obtenues pour toutes les voyelles en AM, AJ et FR en fonction de la situation de production, dans les environnements /b d k/ (sous a) et /b d d^ɣ k/ (sous b). (ns = non significatif ; * = p<0,05 ; ** = p<0,01 ; *** = p<0,001)

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.

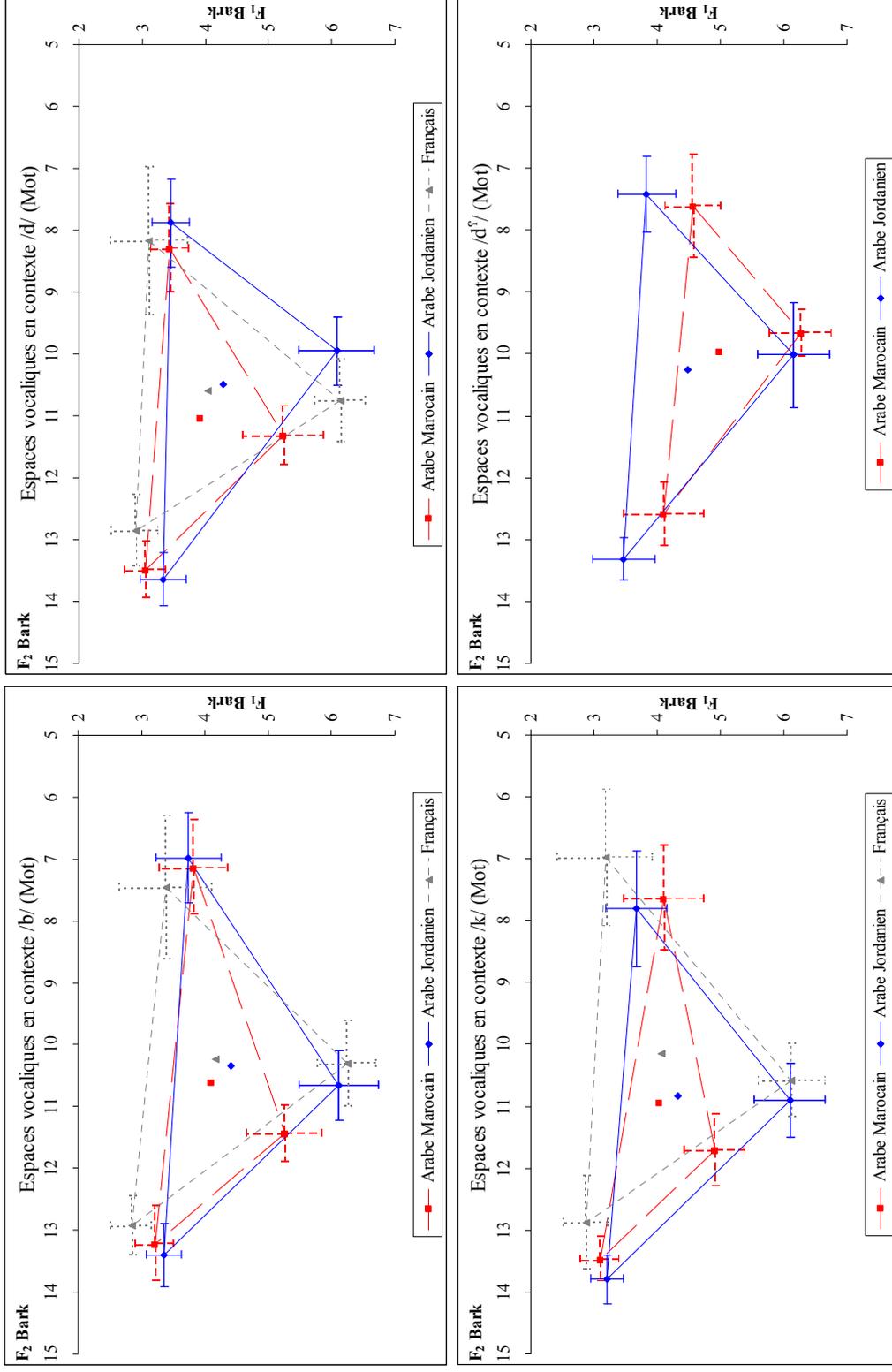


Figure 34 : Dispersion des trois timbres /i a u/ en AM, AJ et FR en fonction de la situation de production Mot et du lieu d'articulation. Les points au milieu de chaque espace vocalique représentent le Barycentre de l'espace pour chaque langue.

RÉSULTATS EN PRODUCTION DES VOYELLES

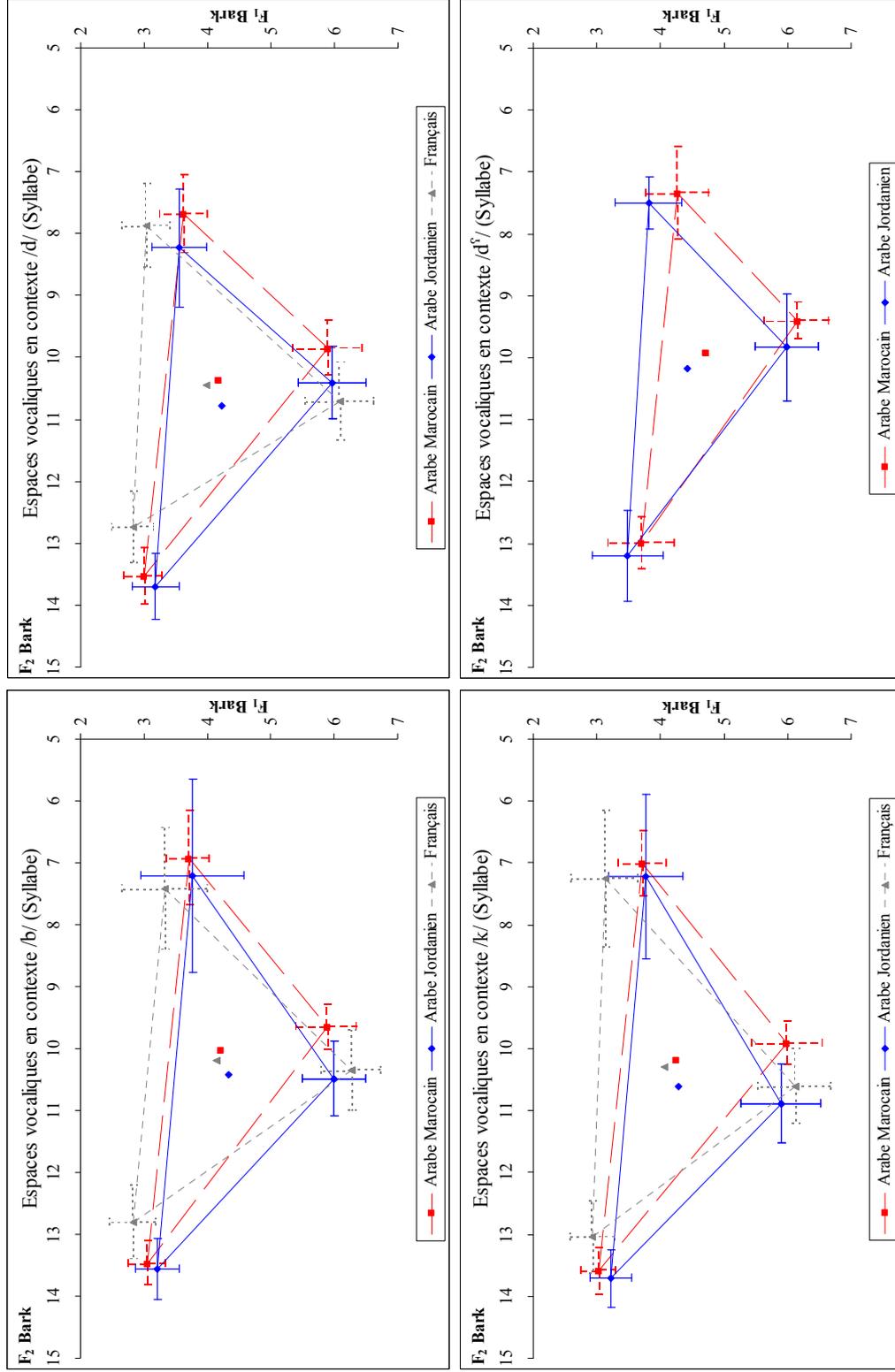


Figure 35 : Dispersion des trois timbres /i u/ en AM, AJ et FR en fonction de la situation de production Syllabe et du lieu d'articulation. Les points au milieu de chaque espace vocalique représentent le Barycentre de l'espace pour chaque langue.

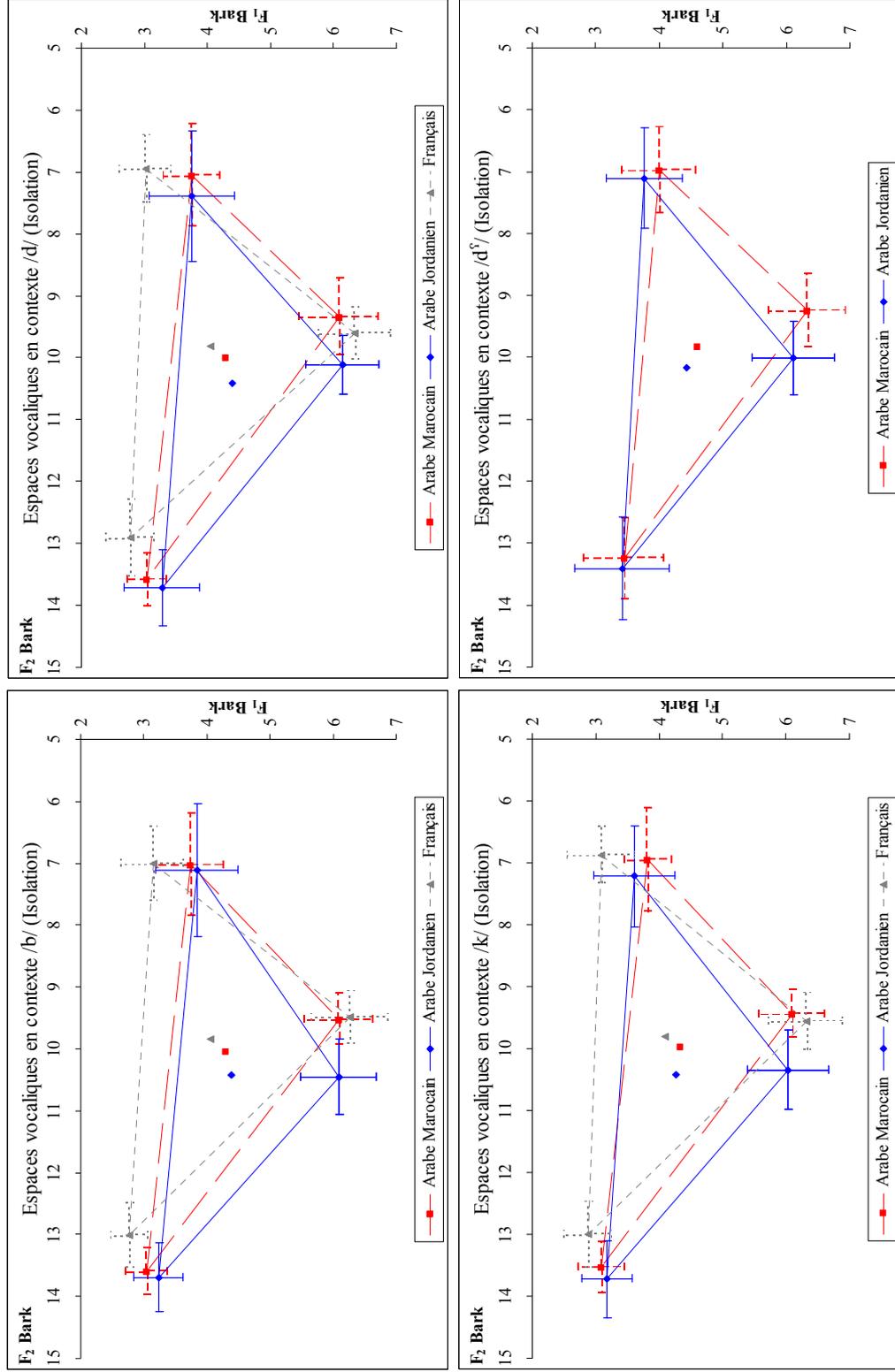


Figure 36 : Dispersion des trois timbres /i u/ en AM, AJ et FR en fonction de la situation de production Isolation et du lieu d'articulation. Les points au milieu de chaque espace vocale représentent le Barycentre de l'espace pour chaque langue.

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.

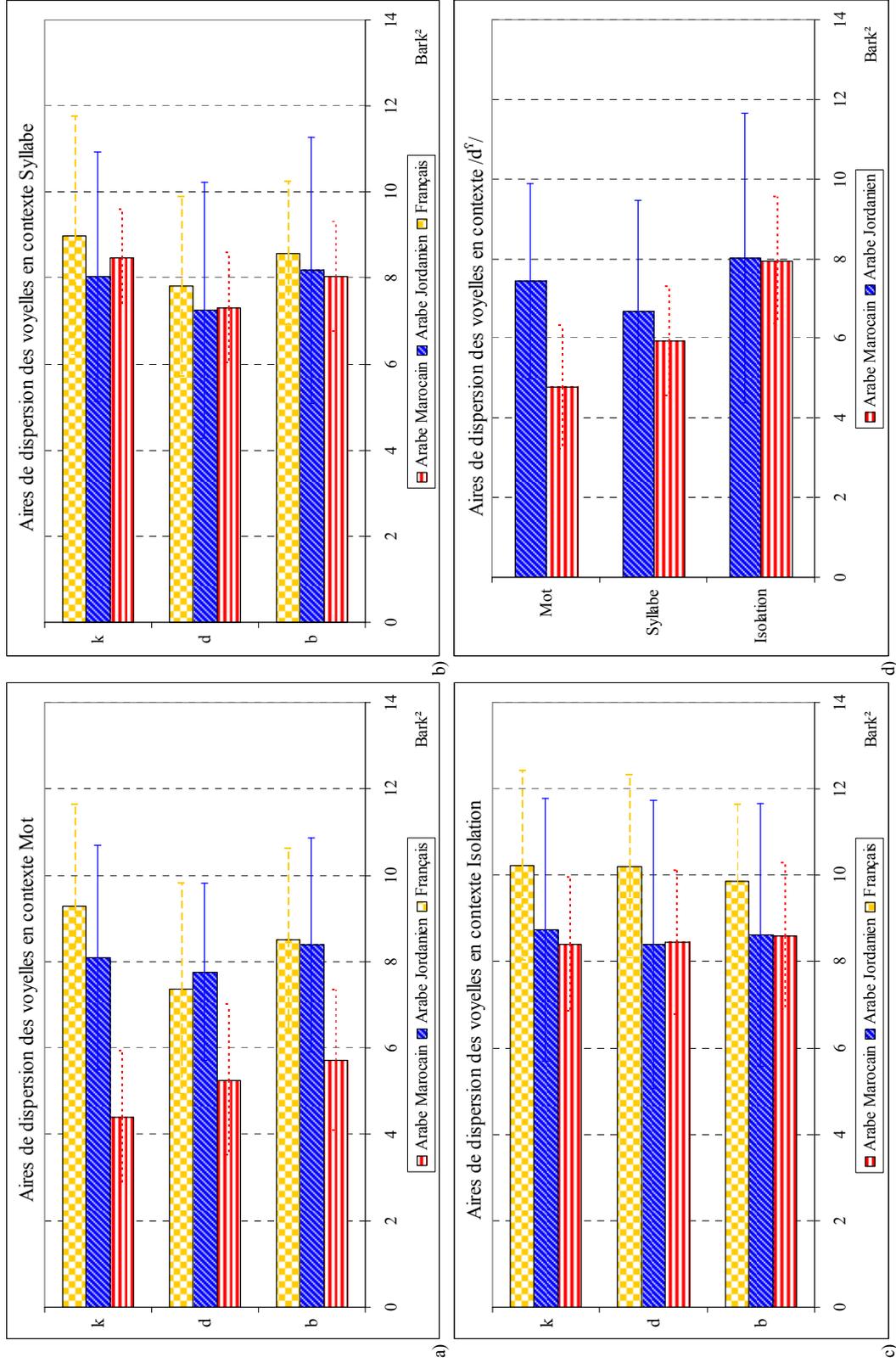


Figure 37 : Aires de dispersion des voyelles en AM, AJ et FR en fonction de la situation de production et du lieu d'articulation de /b d k/. Les aires en contexte /d^h/ en AM et AJ sont présentées sous (d).

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 ≈ 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.

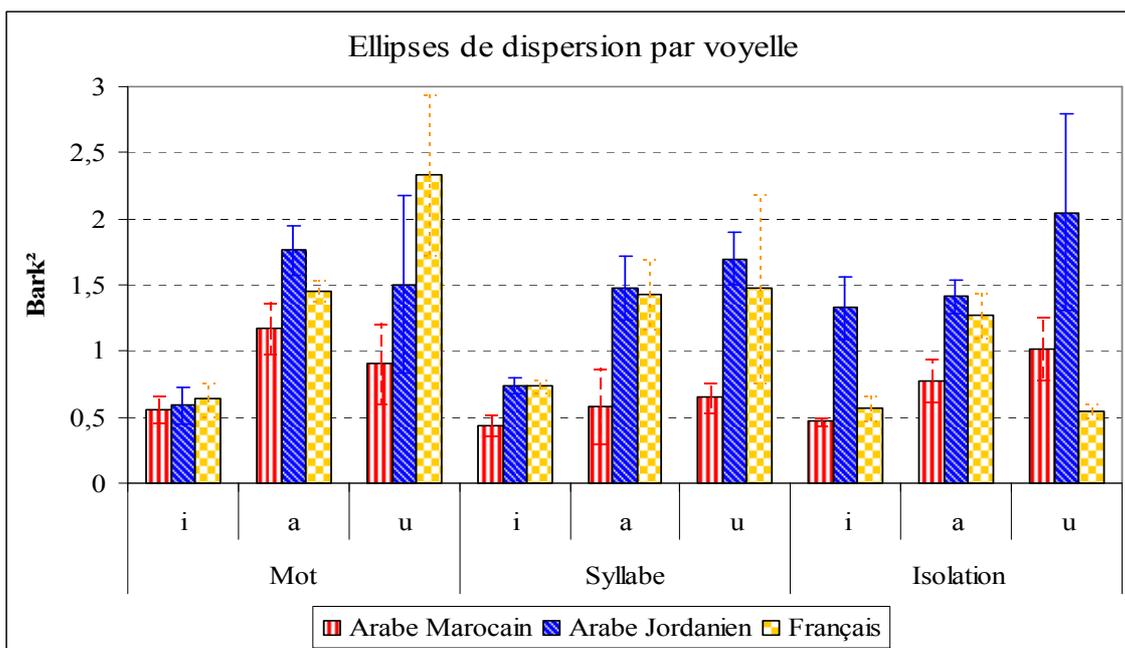


Figure 38 : Ellipses de dispersion des voyelles /i a u/ en AM, AJ et FR en fonction de la situation de production.

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.

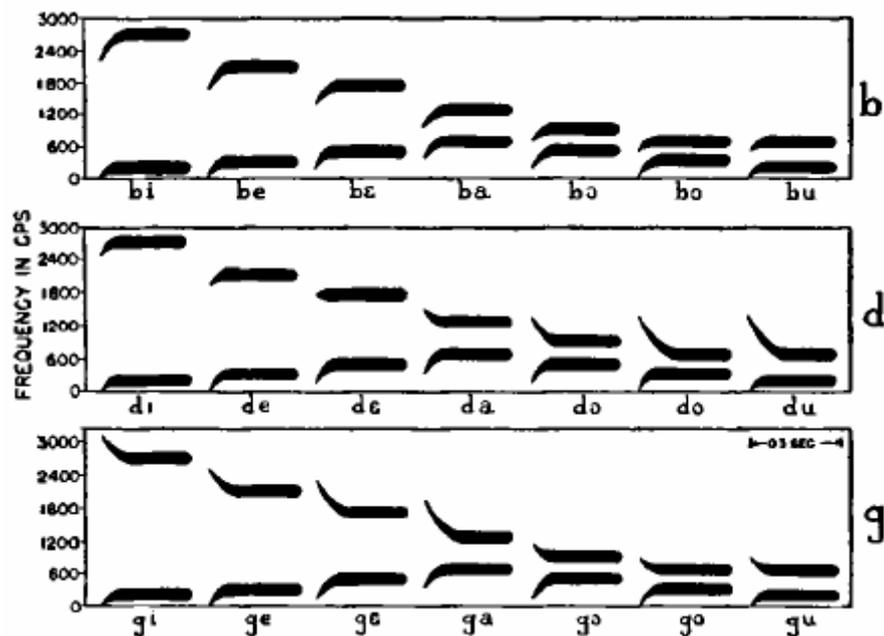


Figure 39 : Direction des transitions de F_1 & de F_2 synthétisées pour différentes voyelles montrant les lieux d'articulation correspondant. (Emprunté à Delattre *et al.*, 1955).

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.

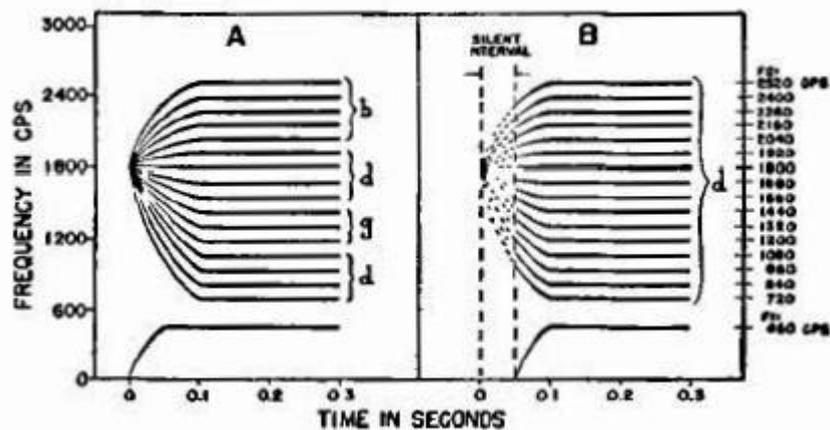


Figure 40 : Stimuli synthétisés avec les deux conditions : transition commençant au « locus » (partie A), et transition commençant à 50 ms après le « locus » (partie B), avec les catégories identifiées par les auditeurs. (Emprunté à Delattre *et al.*, 1955).

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.

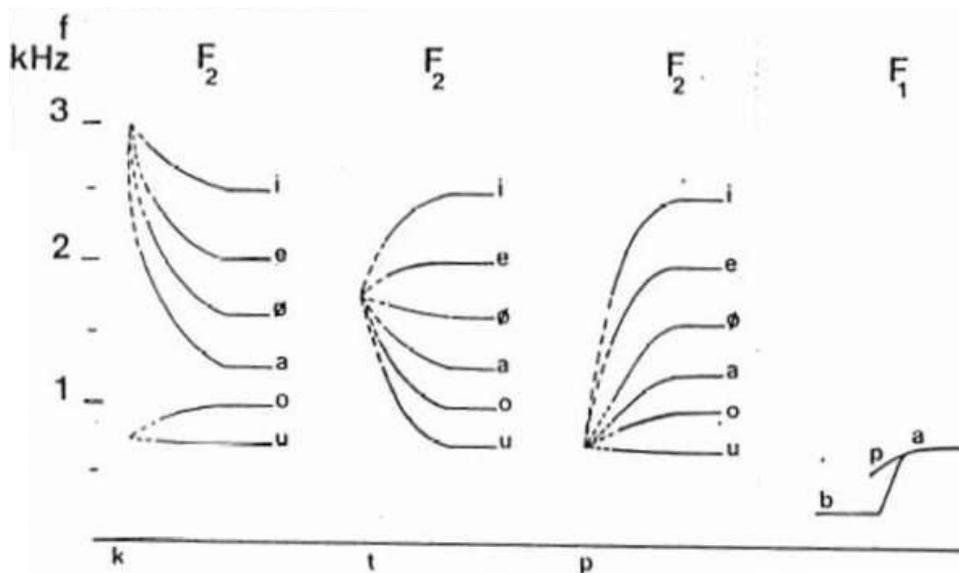


Figure 41 : Directions des transitions formantiques dans des syllabe CV avec la position du « locus virtuel » de F_2 en fonction des consonnes /p t k/ et des voyelles /i e ø a o u/. (Emprunté à Landercy & Renard, 1977).

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 » = $b / (1 - m)$).

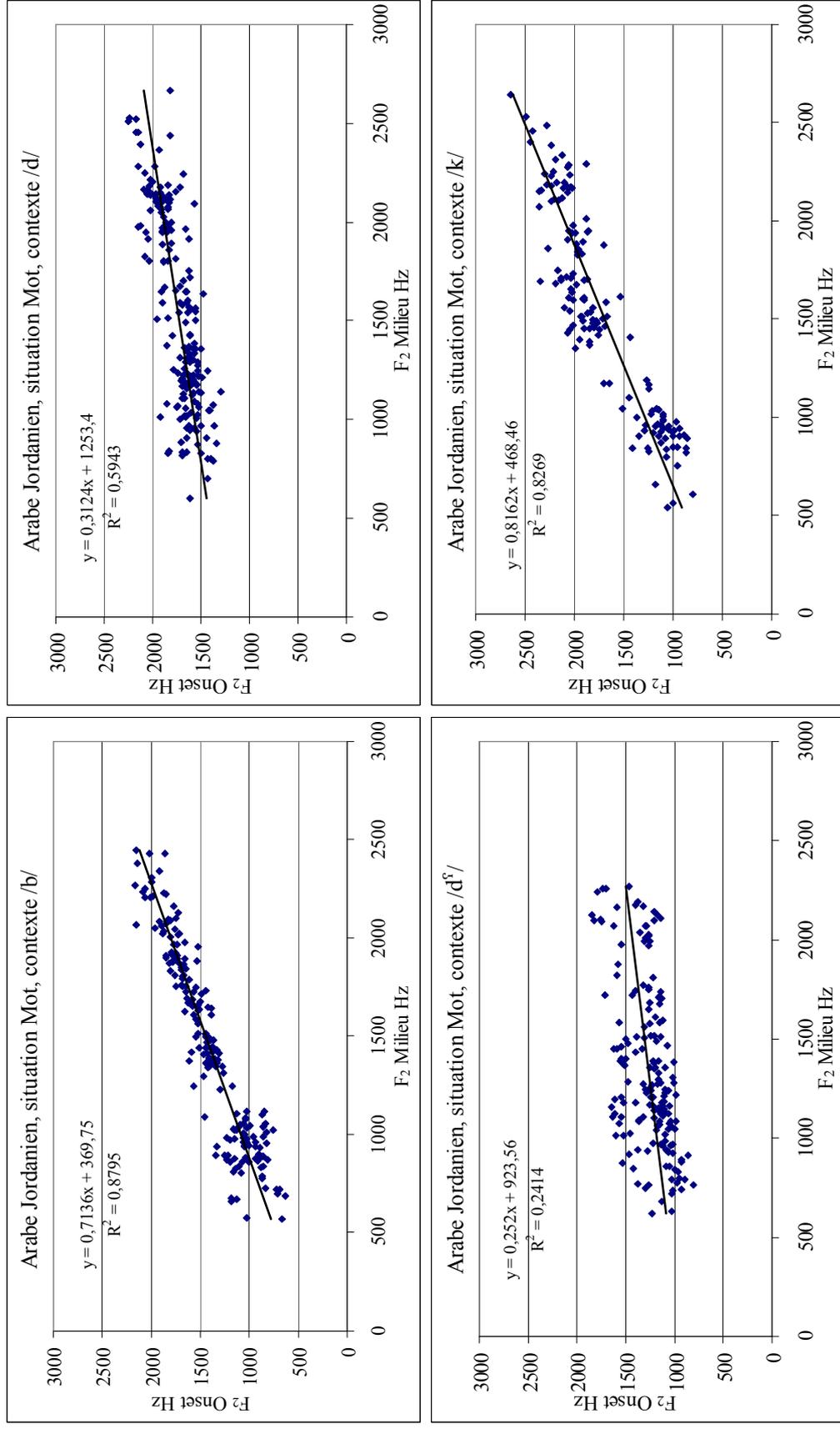


Figure 42 : Représentation en nuage de points de la relation entre F_{2Onset} et $F_{2Milieu}$ et la droite de régression pour les consonnes /b d^v k/ de la coarticulation CV, avec les coefficients de l'équation de locus. Les points indiquent des voyelles.

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.

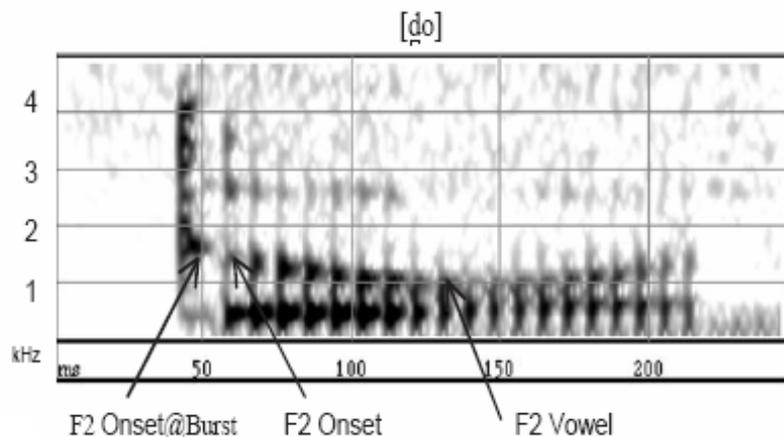


Figure 43 : Schéma présentant les méthodes d'extraction de l'**onset** : a) @Burst et b) à la première pulsation après le relâchement de la voyelle. (Emprunté à Modarresi *et al.*, 2005).

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.

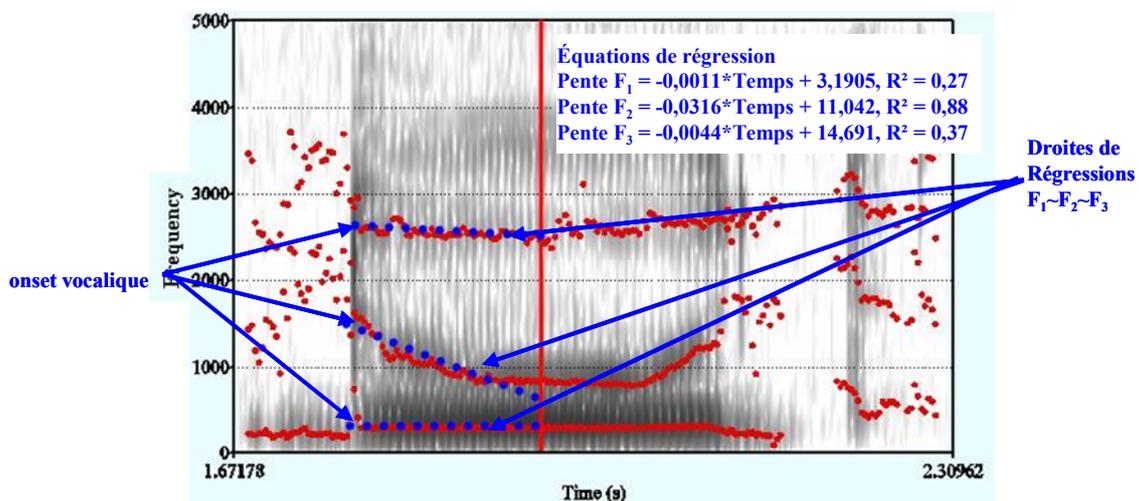


Figure 44 : Schéma représentant le calcul de régression linéaire pour la voyelle /u:/ comme réalisée dans le mot /du:d/ « vers de terre (pl.) » par le locuteur 09MAJ, avec la position de l'onset vocalique, ainsi que les droites et les équations de régression linéaire.

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 **pente 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).

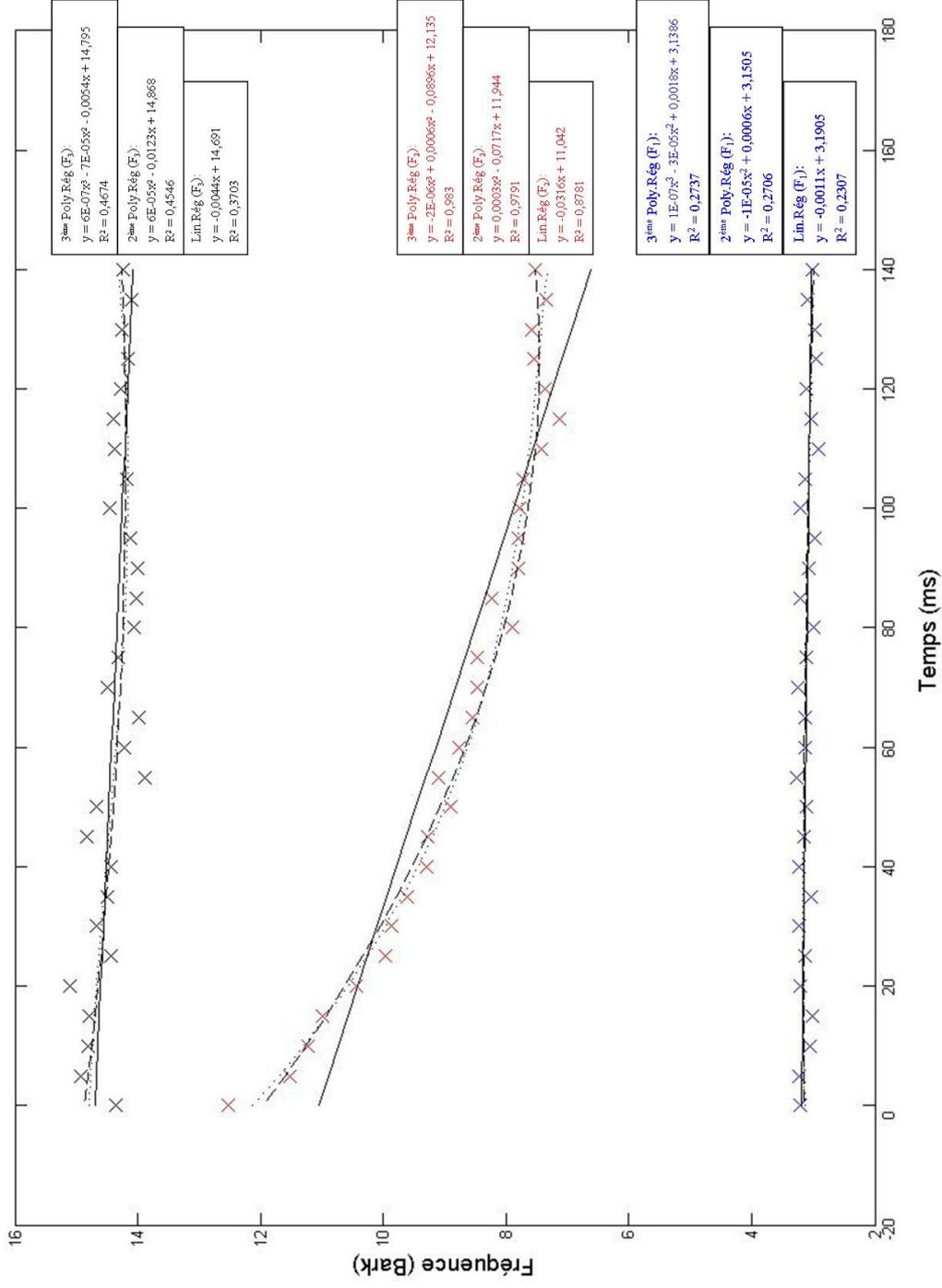


Figure 45 : Droites de régression linéaires (en lignes continues), polynômes de 2^{ème} ordre (en lignes discontinues) et polynômes de 3^{ème} ordre (en lignes pointillées) de F₁, F₂ & F₃ de la voyelle /u:/ comme réalisée dans le mot /dur:/ « vers de terre (pl.) » par le locuteur 09MAJ, ainsi que les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

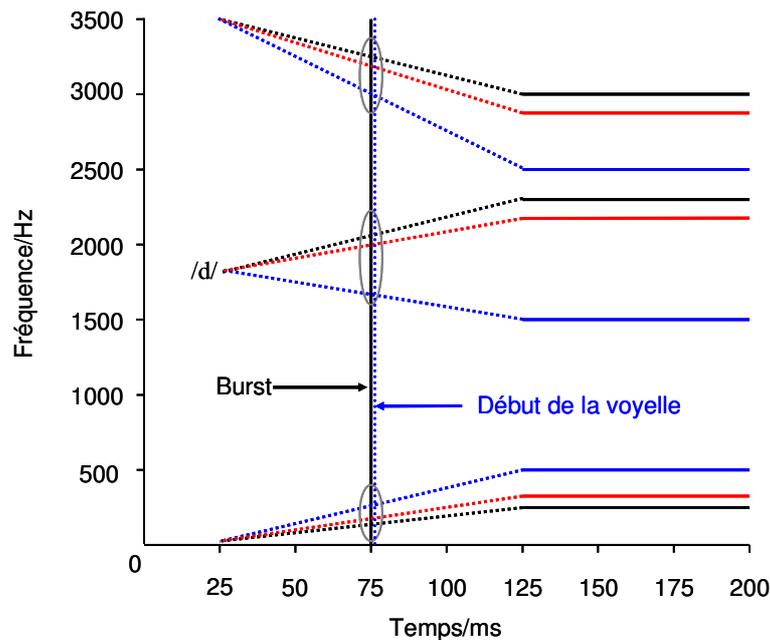


Figure 46 : Syllabe /di/ schématisée en trois conditions : voyelle /i/ pleine non réduite (en ligne noire) ; voyelle /i/ pleine légèrement réduite (en ligne rouge) et voyelle /i/ réduite en /ə/ (en ligne bleue). Les cercles entourent les **onsets** vocaliques extraits juste après le début de la voyelle.

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 ($F(3, 4768)=35,07$; $p<0,001$).

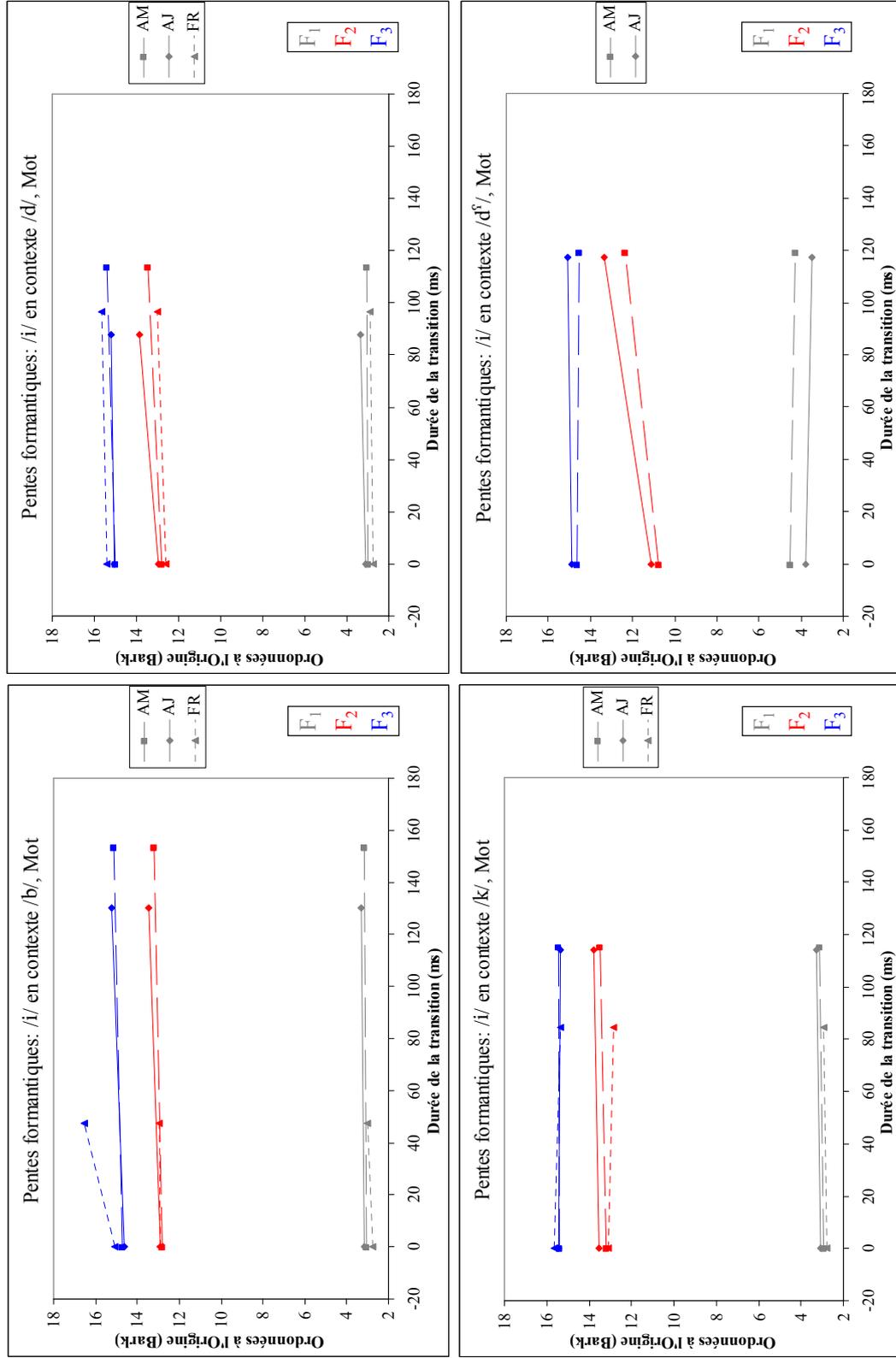


Figure 47 : Pentes formantiques de F₁, F₂ & F₃ sans normalisation du temps, pour la voyelle /i/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

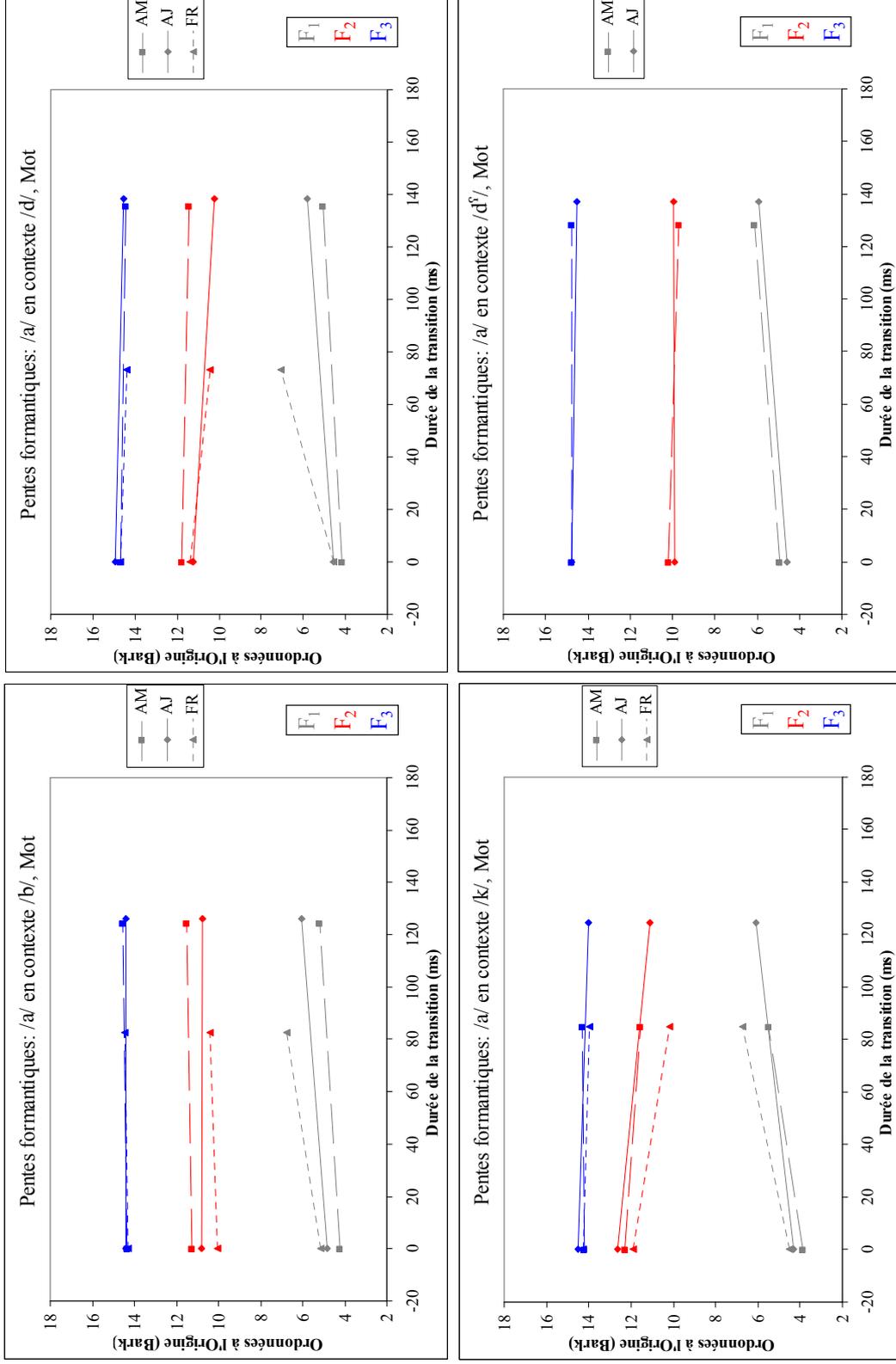


Figure 48 : Pentas formantiques de F_1 , F_2 & F_3 sans normalisation du temps, pour la voyelle /a/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

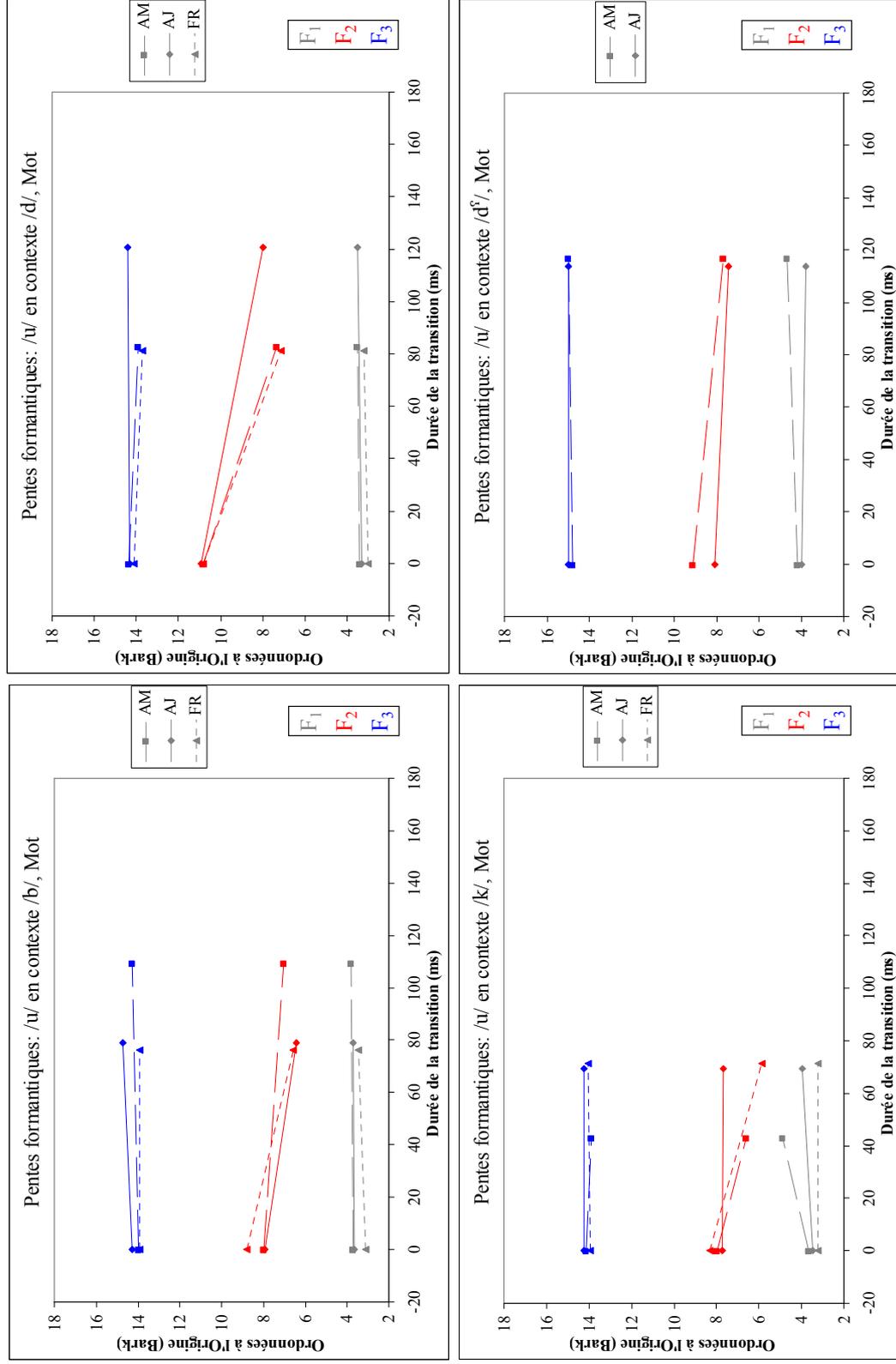


Figure 49 : Pentas formantiques de F_1 , F_2 & F_3 sans normalisation du temps, pour la voyelle /u/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

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).

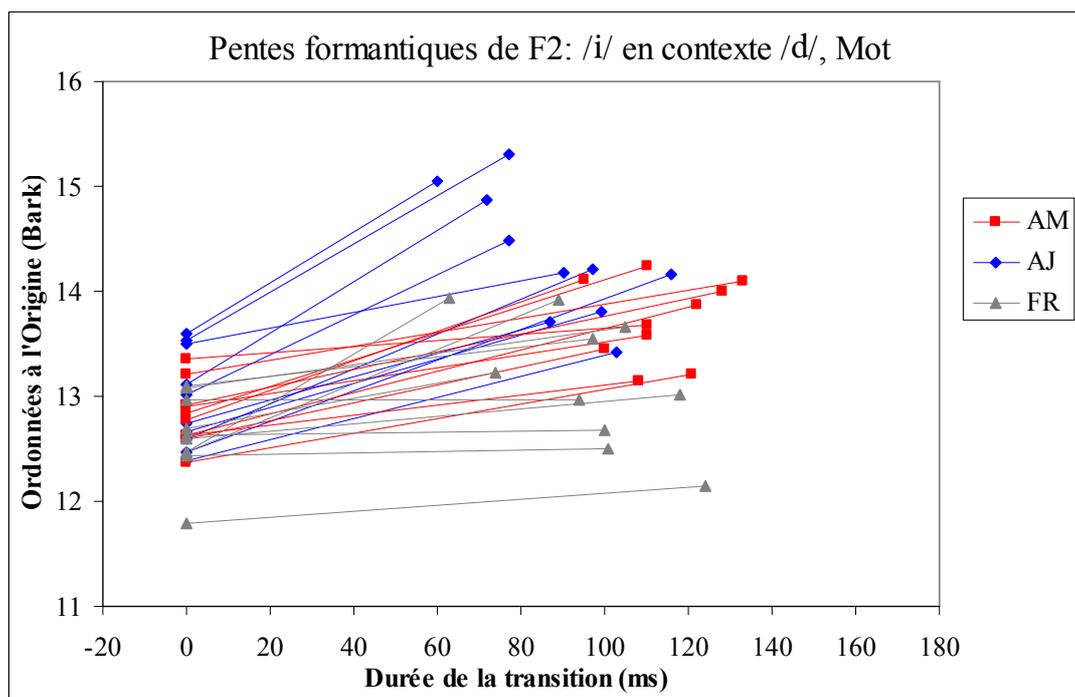


Figure 50 : Pentas formantiques individuelles moyennes, non normalisées, de F_2 de la voyelle /i/ produite dans la situation Mot dans l'environnement /d/ dans les trois langues.

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 F_2 , 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 **pente 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 **pente 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 pentés 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.

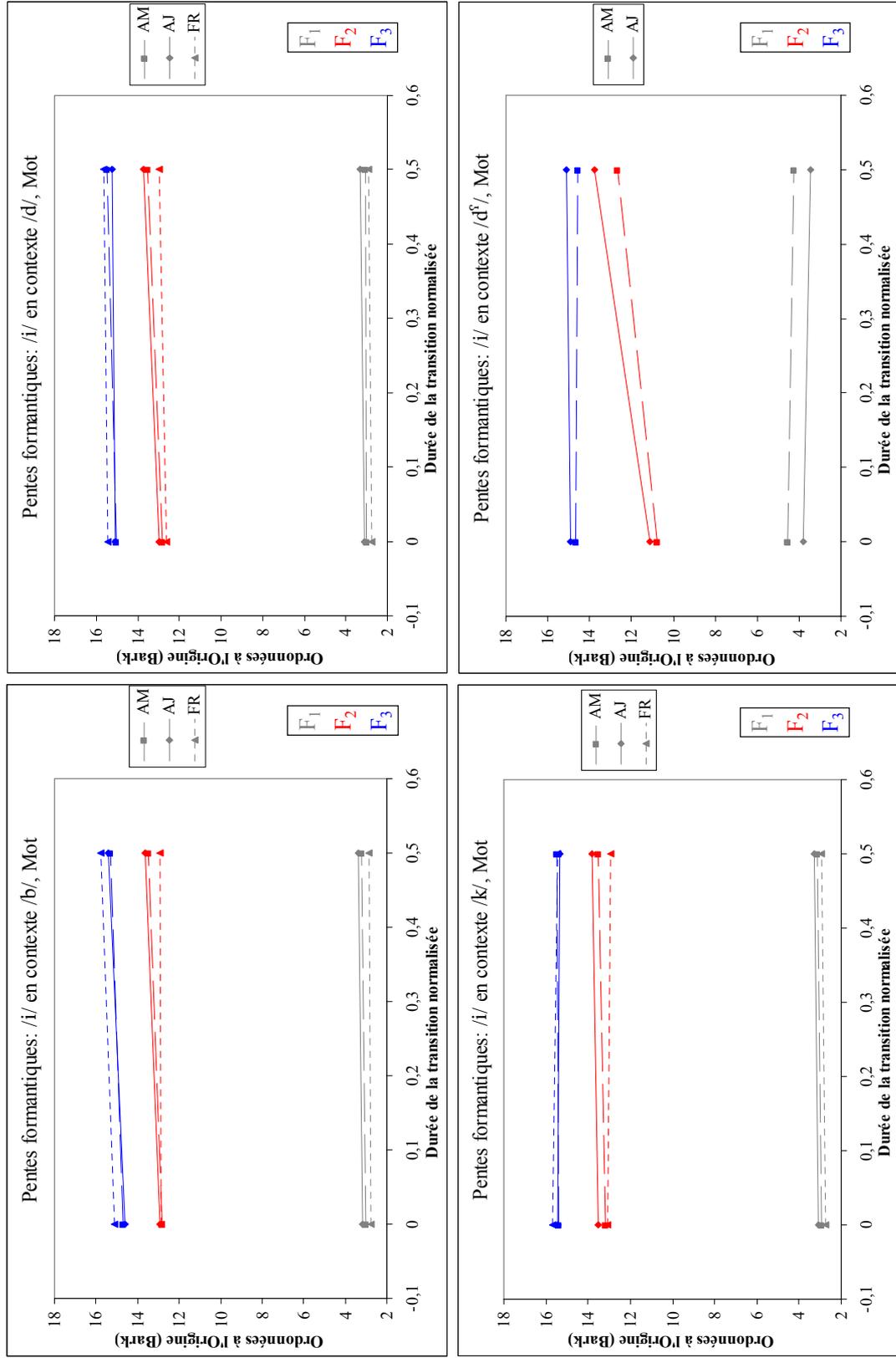


Figure 51 : Pentas formantiques de F_1 , F_2 & F_3 avec normalisation du temps, pour la voyelle /i/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

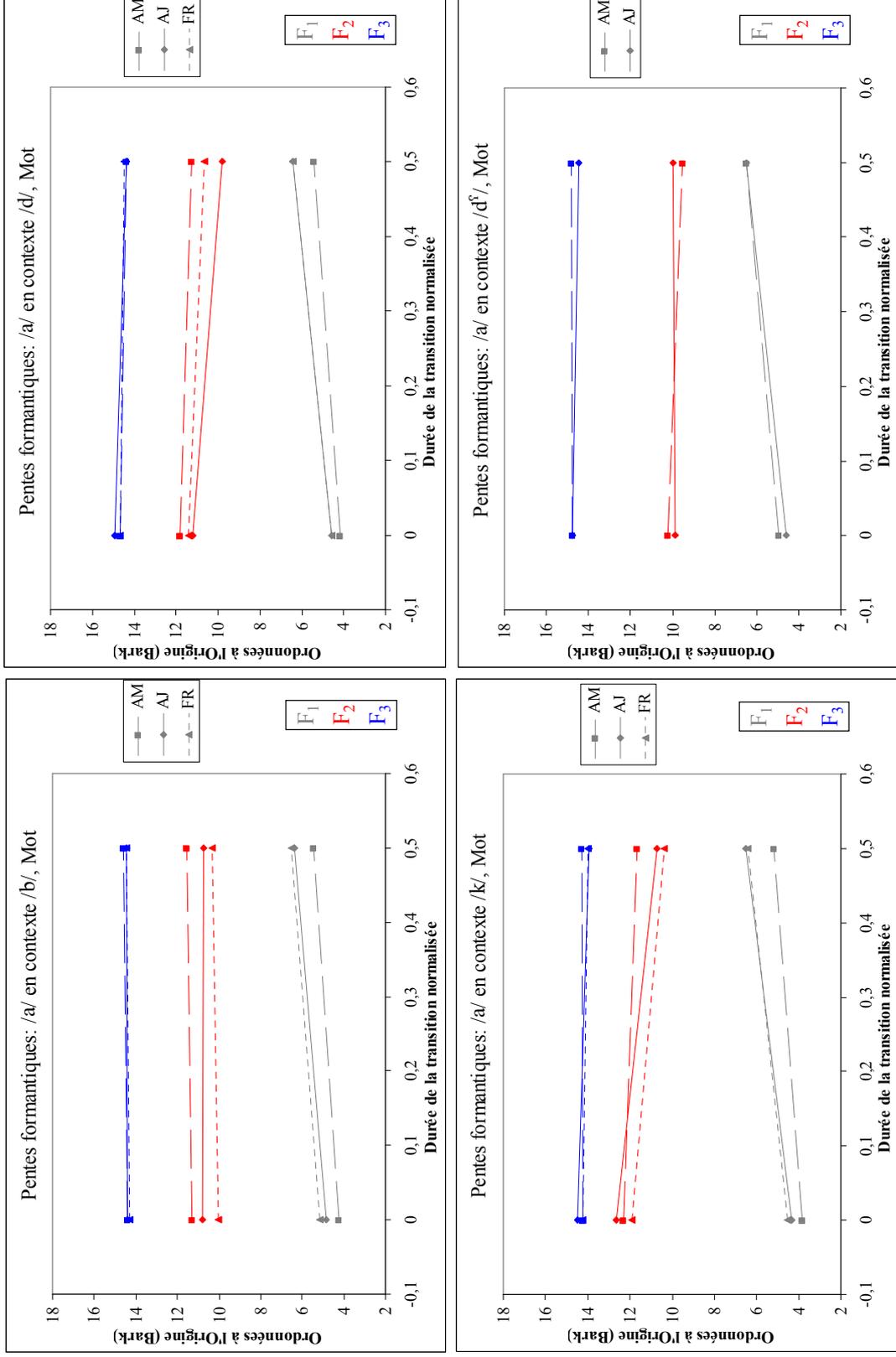


Figure 52 : Pentas formantiques de F_1 , F_2 & F_3 avec normalisation du temps, pour la voyelle /a/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

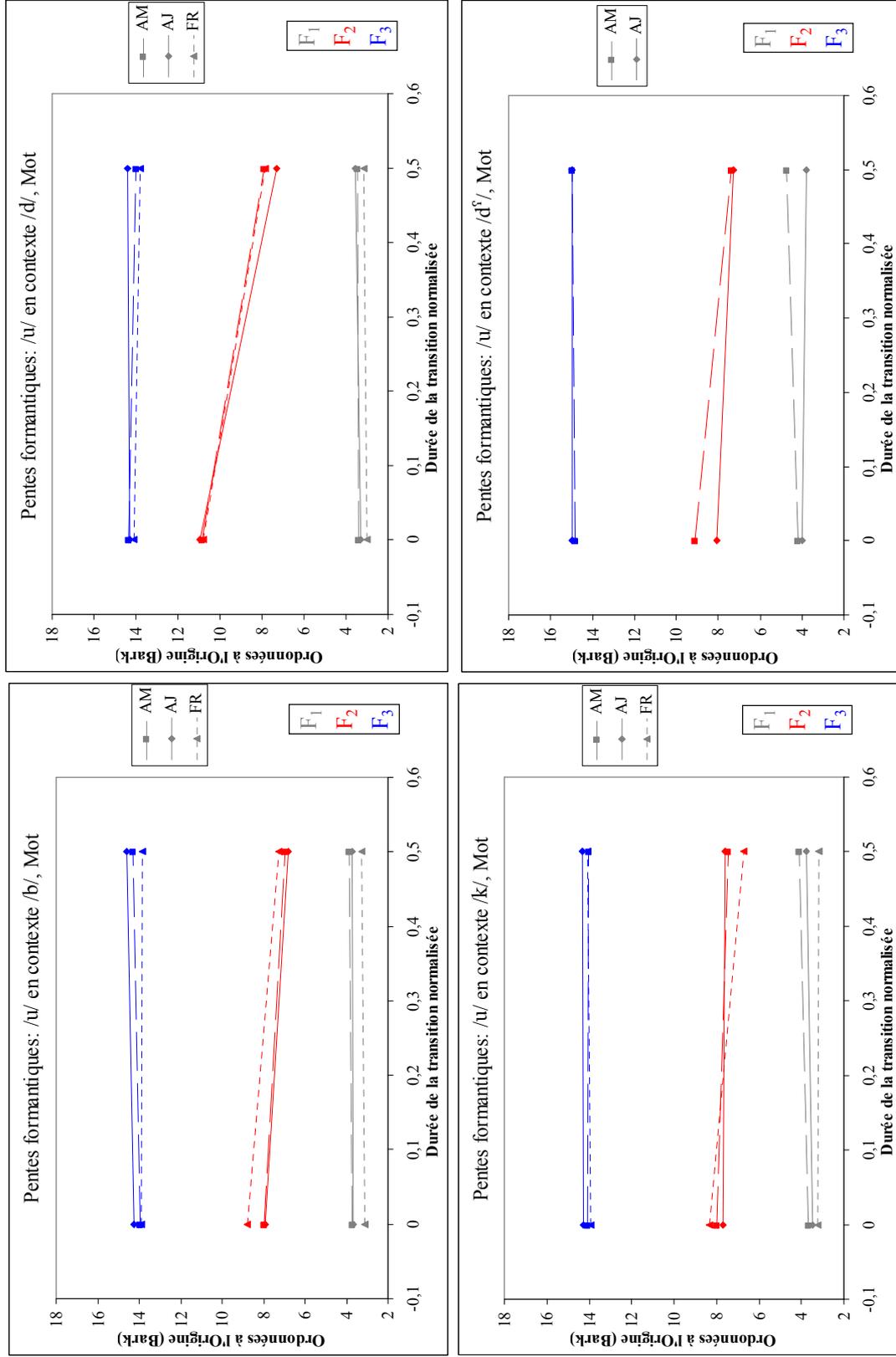


Figure 53 : Pentas formantiques de F₁, F₂ & F₃ avec normalisation du temps, pour la voyelle /u/ dans les trois langues en fonction du lieu d'articulation en situation Mot.

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 **pente 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 **pente 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 **pente 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.

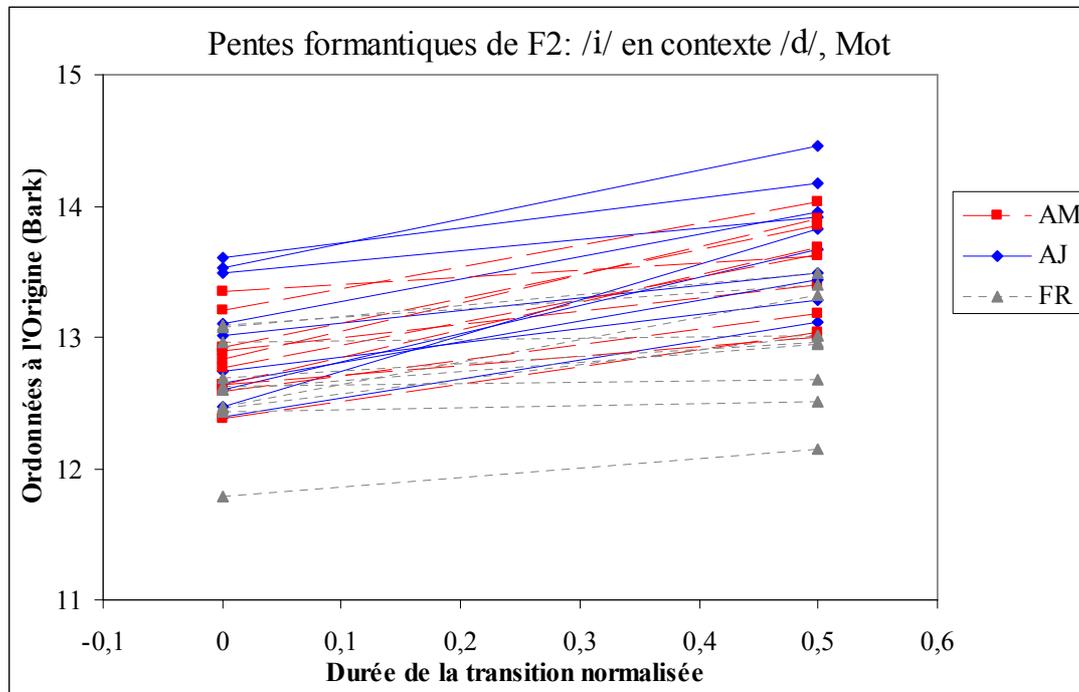


Figure 54 : Pentas formantiques individuelles moyennes, normalisées, de F₂ de la voyelle /i/ produite en situation de production Mot dans l'environnement /d/ dans les trois langues.

En examinant plus en détail les **pentas 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 **pentas 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 dépend 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 ; $p < 0,001$).

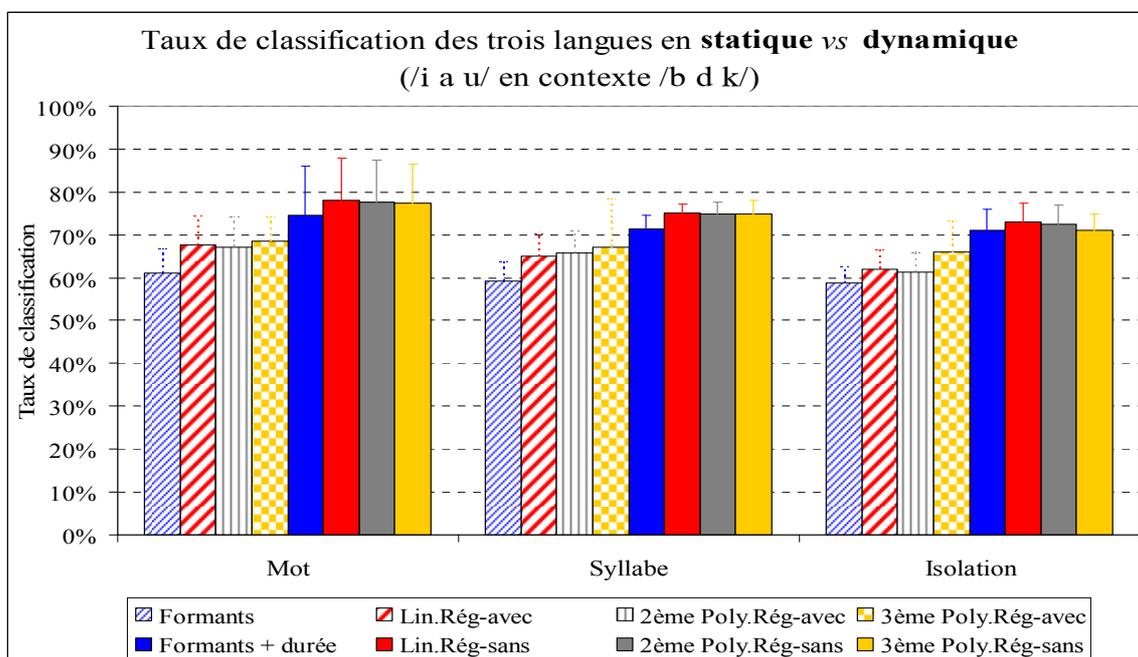


Figure 55 : Taux de classification correcte moyens des voyelles /i a u/ produites dans les contextes /b d k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur), dans les trois langues. Lin.Rég = régression linéaire, Poly.Rég = régression 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).

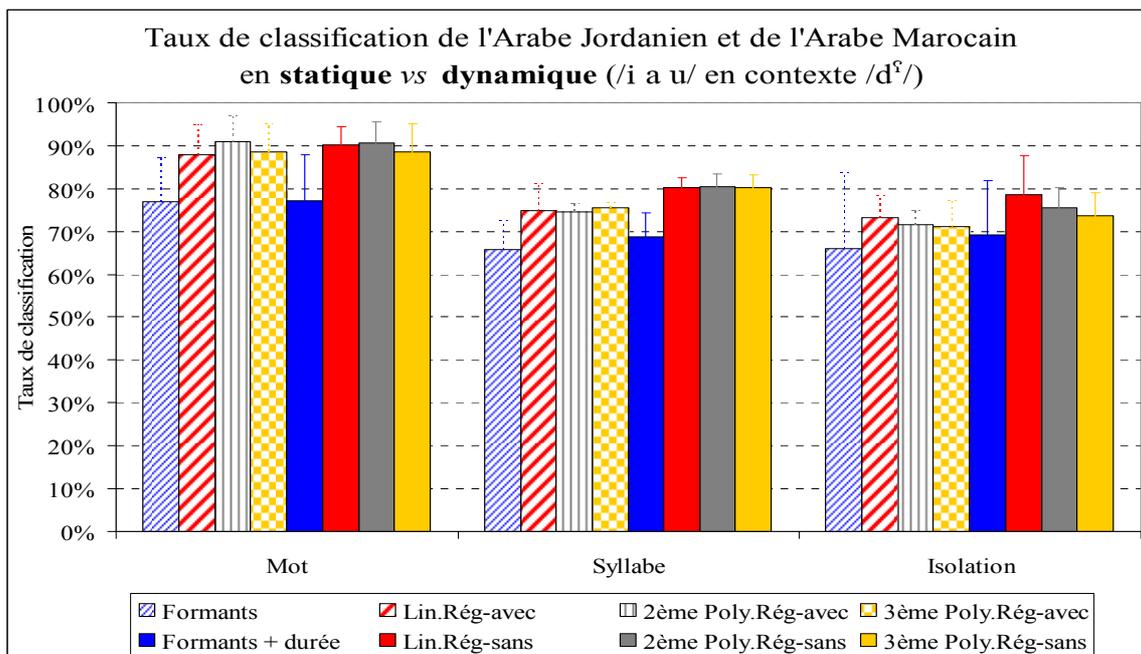


Figure 56 : Taux de classification correcte moyens des voyelles /i a u/ produites dans le contexte /d^ʕ/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur) en arabe jordanien et arabe marocain. Lin.Rég = régression linéaire, Poly.Rég = 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.

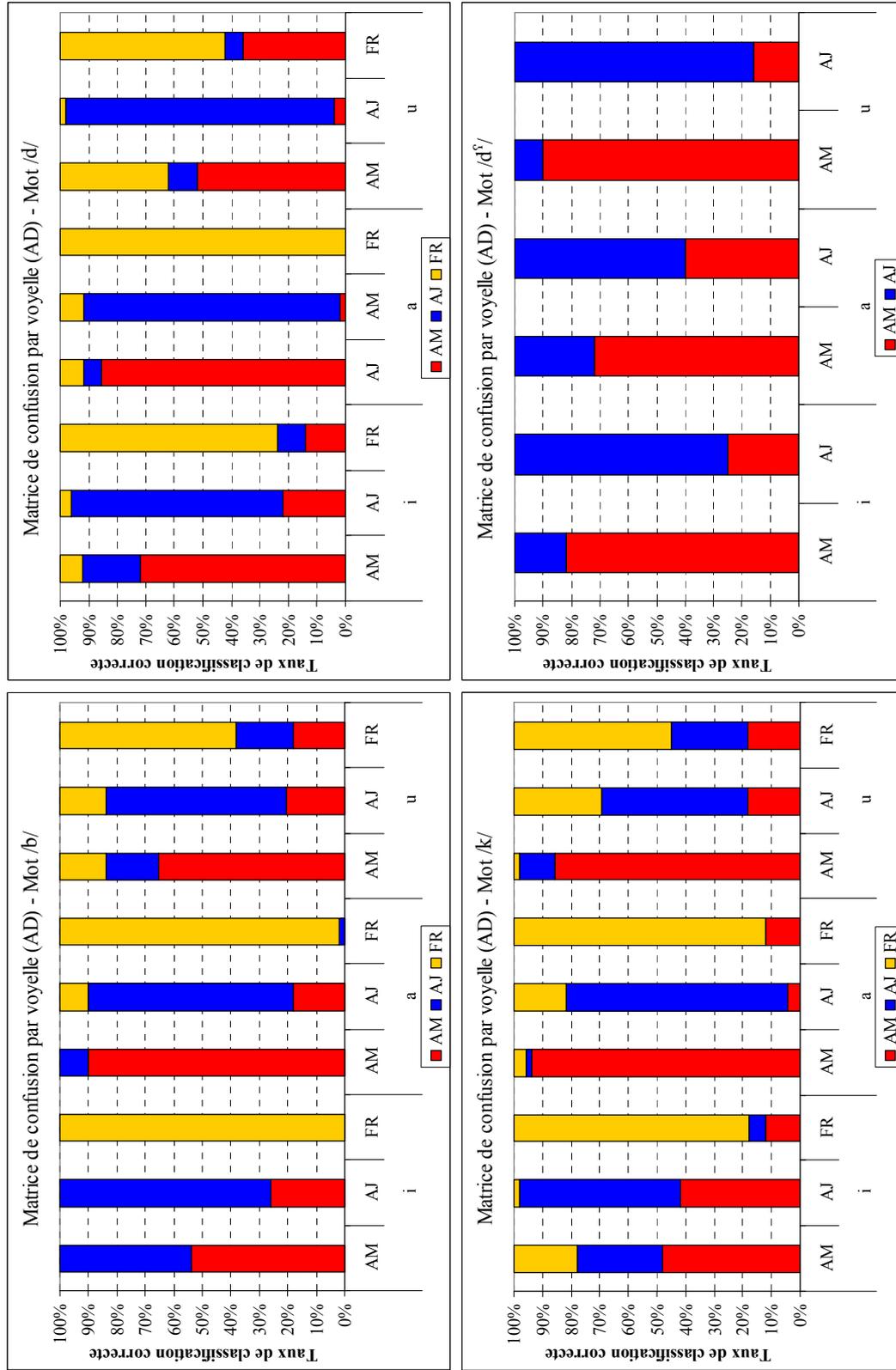


Figure 57 : Matrices de confusion de classification de chaque voyelle /i a u/ en **statique** avec la durée de la transition (AD) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d d^f k/ (/d^f/ entre AM & AJ).

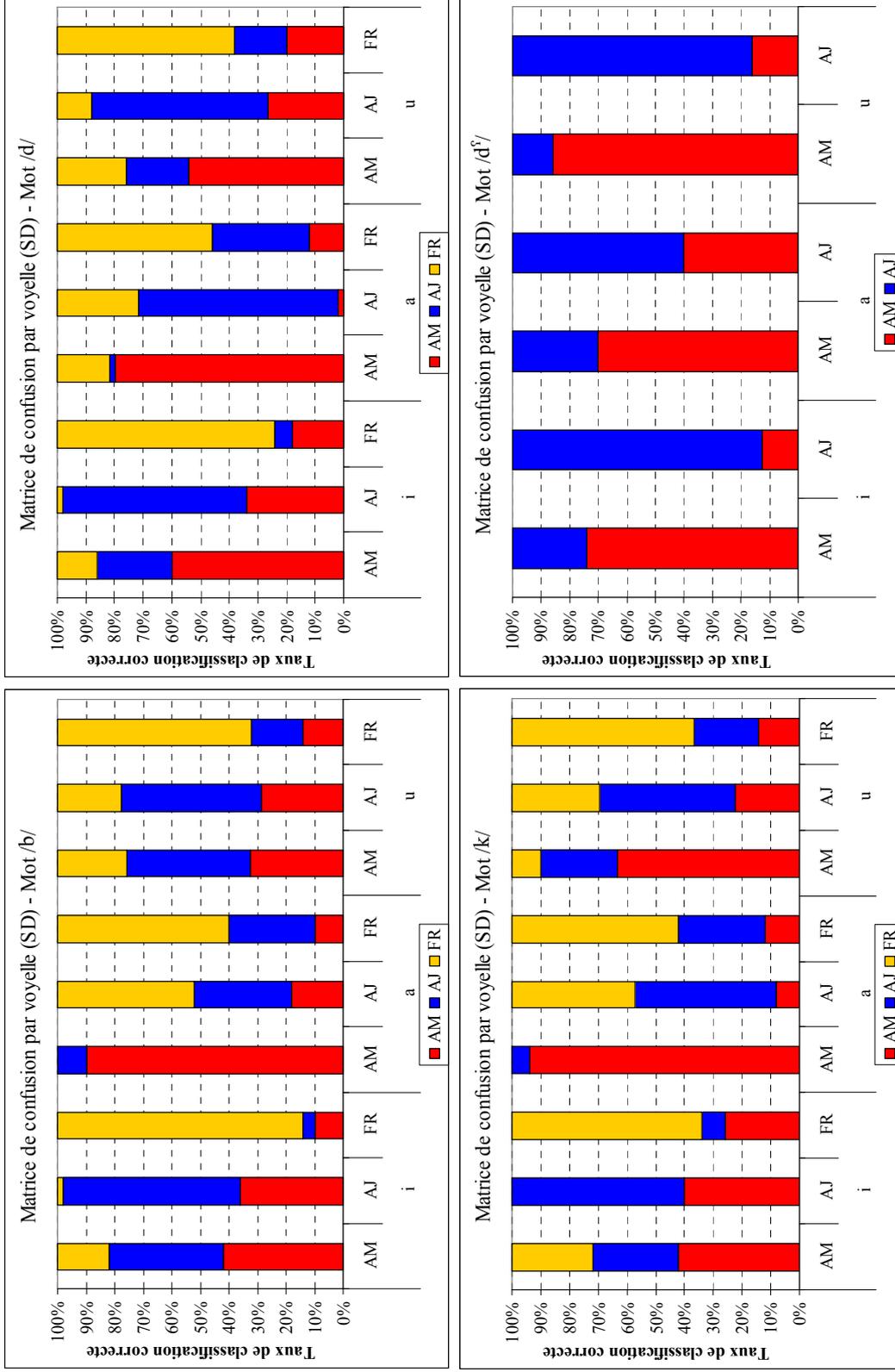


Figure 58 : Matrices de confusion de classification de chaque voyelle /i a u/ en **statique** sans la durée de la transition (SD) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b dʰ k/ (/dʰ/ entre AM & AJ).

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).

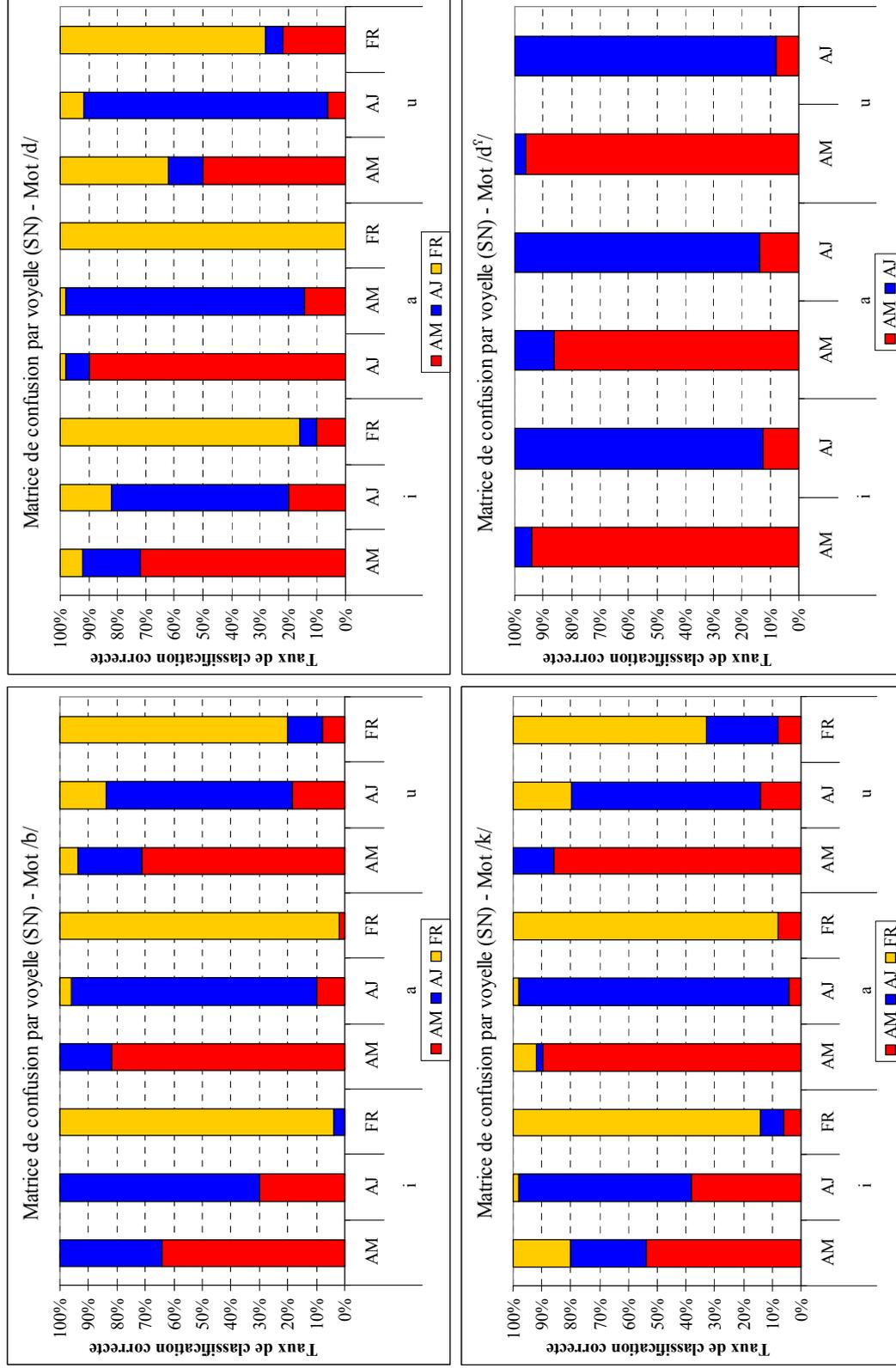


Figure 59 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression linéaire) sans normalisation du temps (SN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d d^h k/ (/d^h/ entre AM & AJ).

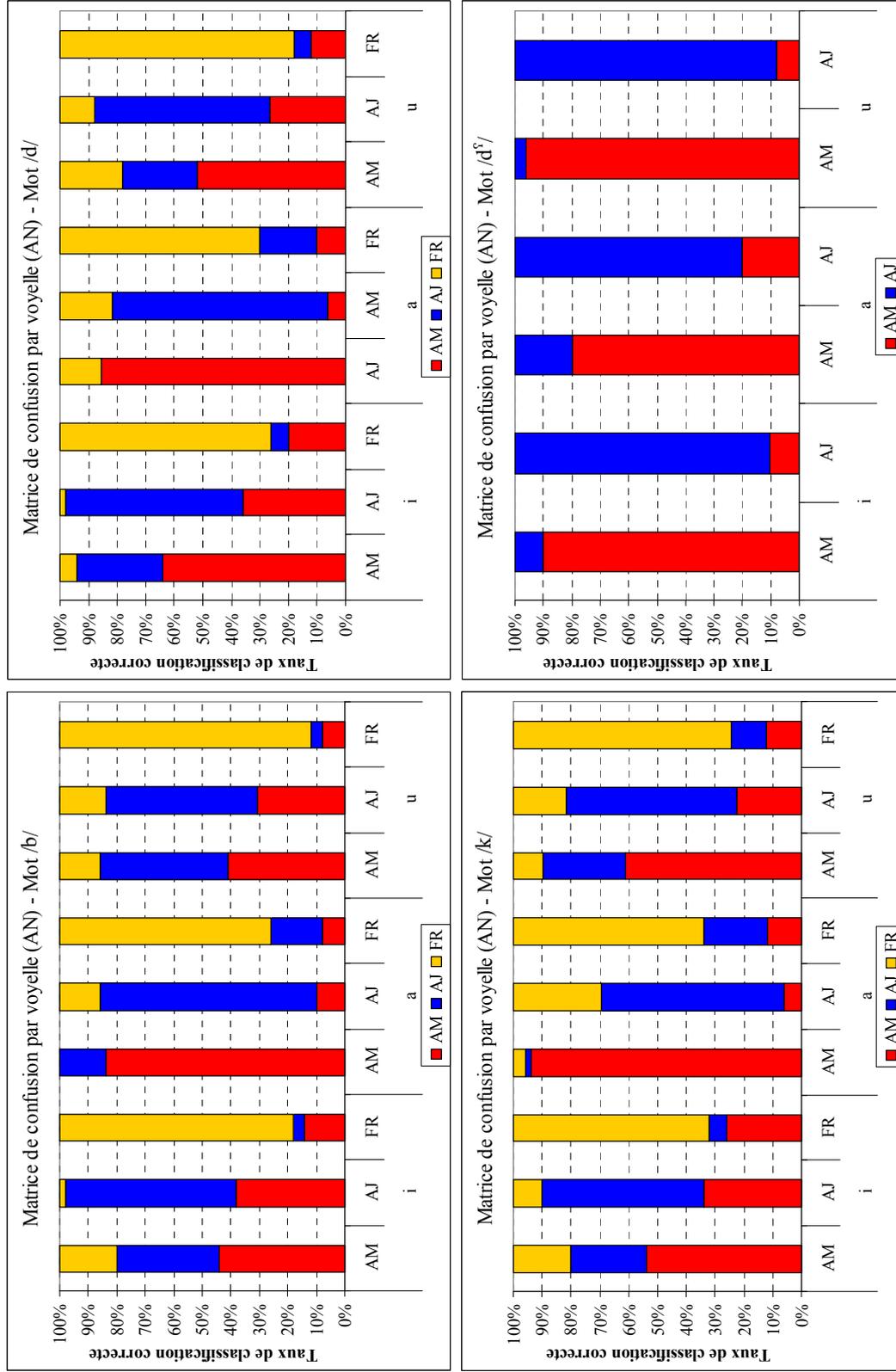


Figure 60 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression linéaire) avec normalisation du temps (AN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d d^f k/ (/d^f k/ (/d^f entre AM & AJ).

Ces résultats montrent que les coefficients de la régression linéaire : les **pentés 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 **pentés 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 **pentés 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.

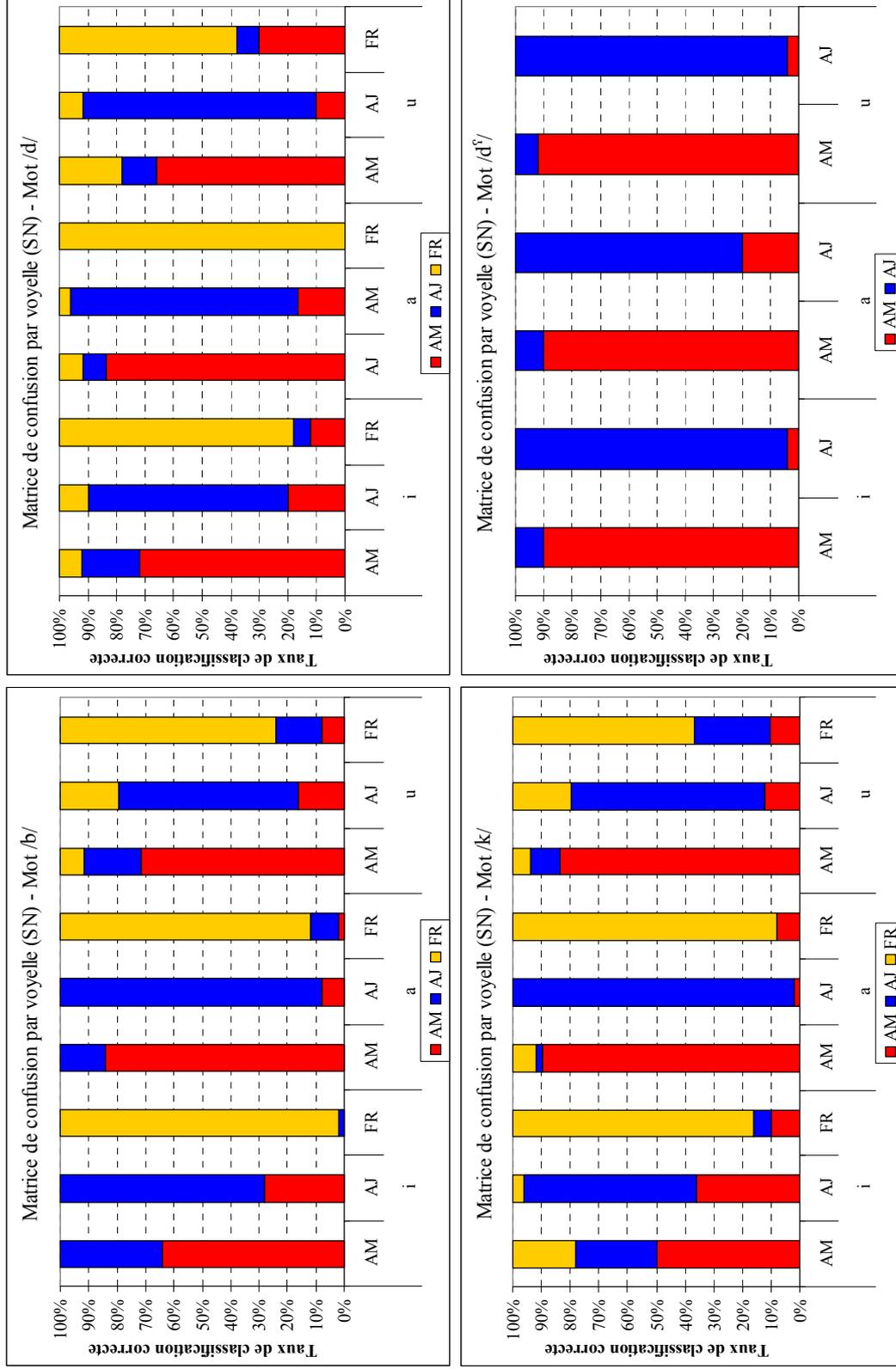


Figure 61 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression polynomiale 2^{ème} ordre) sans normalisation du temps (SN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f entre AM &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

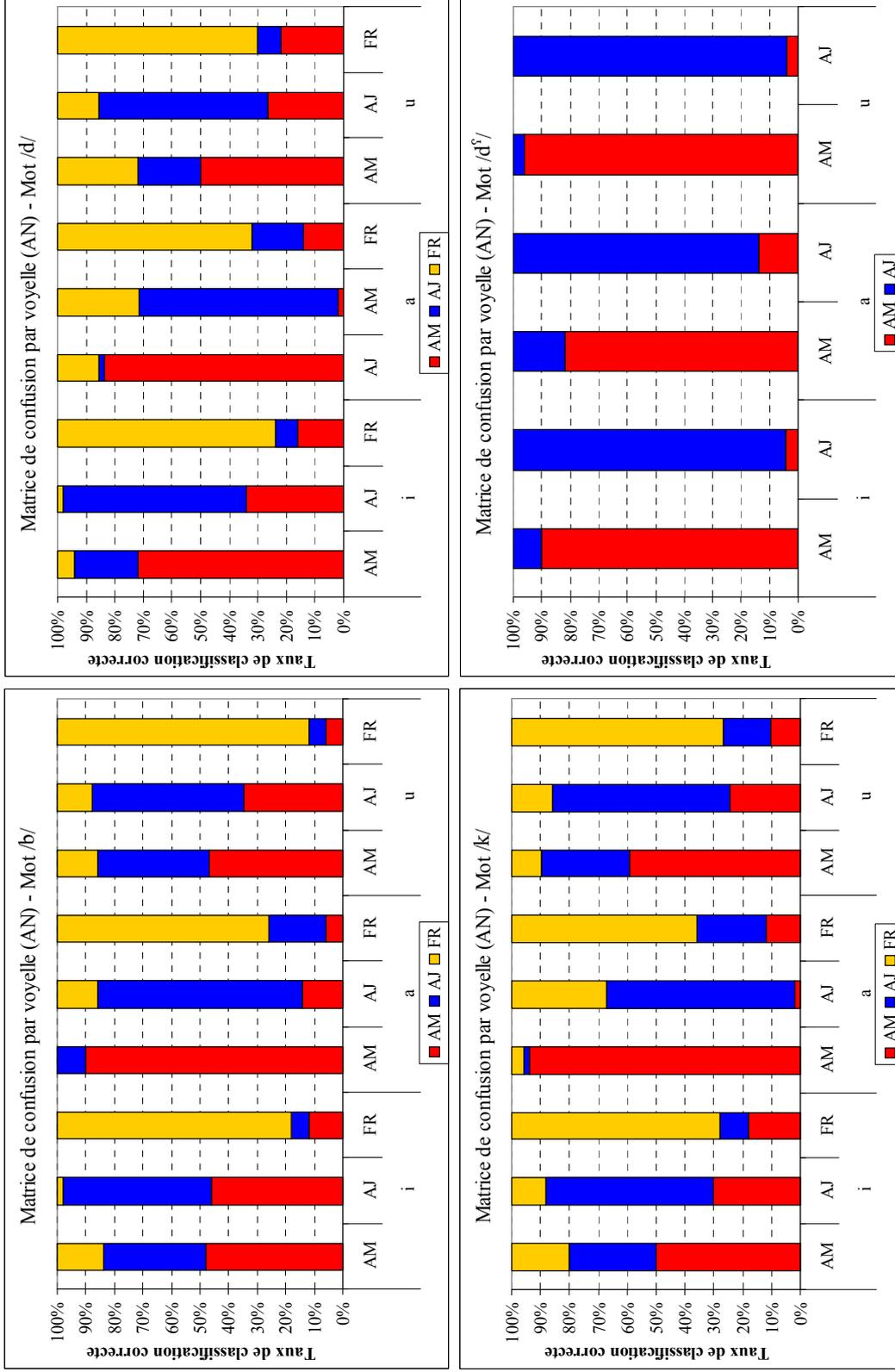


Figure 62 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression polynomiale 2^{ème} ordre) avec normalisation du temps (AN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f/ entre AM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

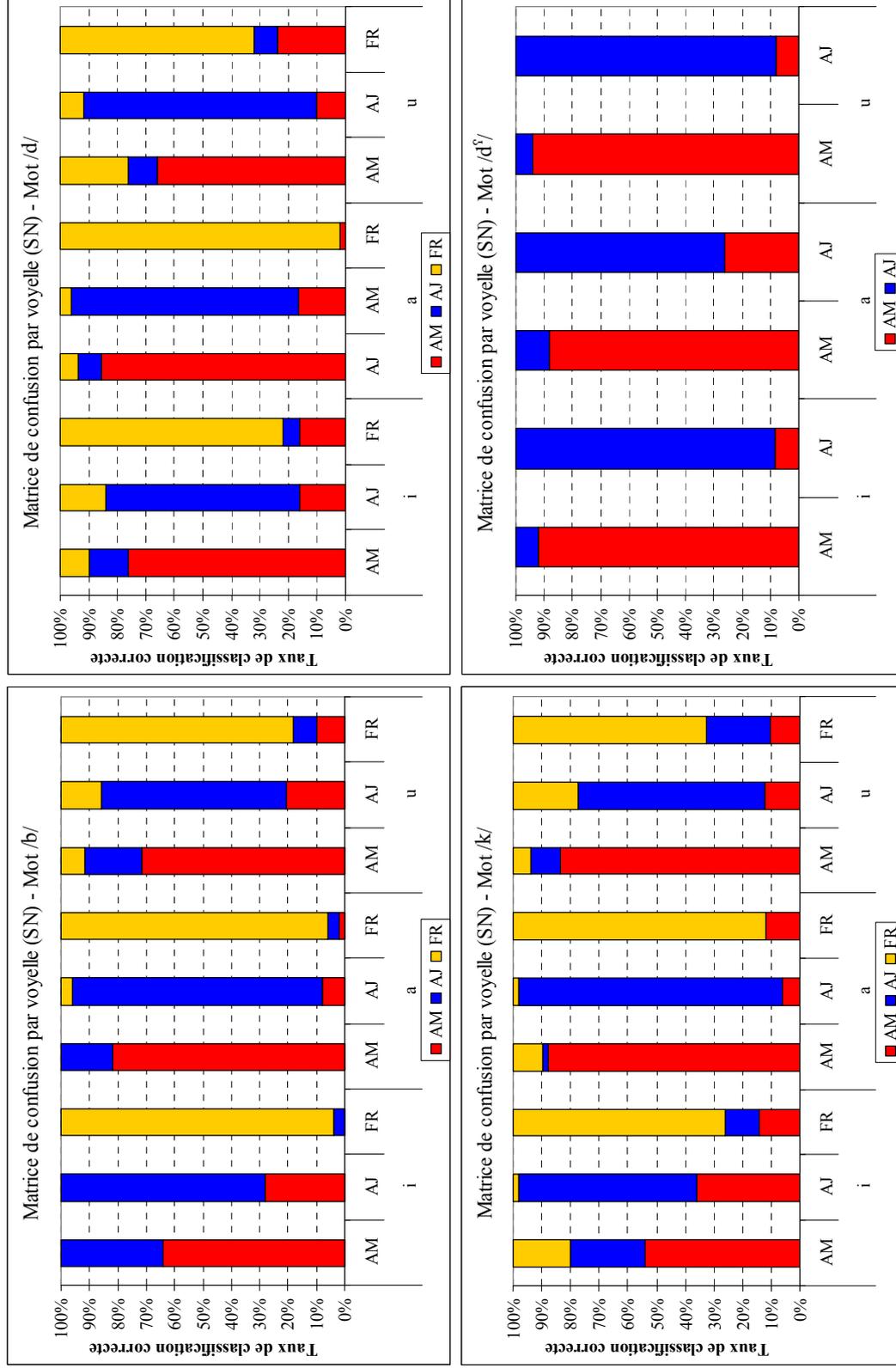


Figure 63 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression polynomiale 3^{ème} ordre) sans normalisation du temps (SN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d̥ k/ (/d̥/ entre AM &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

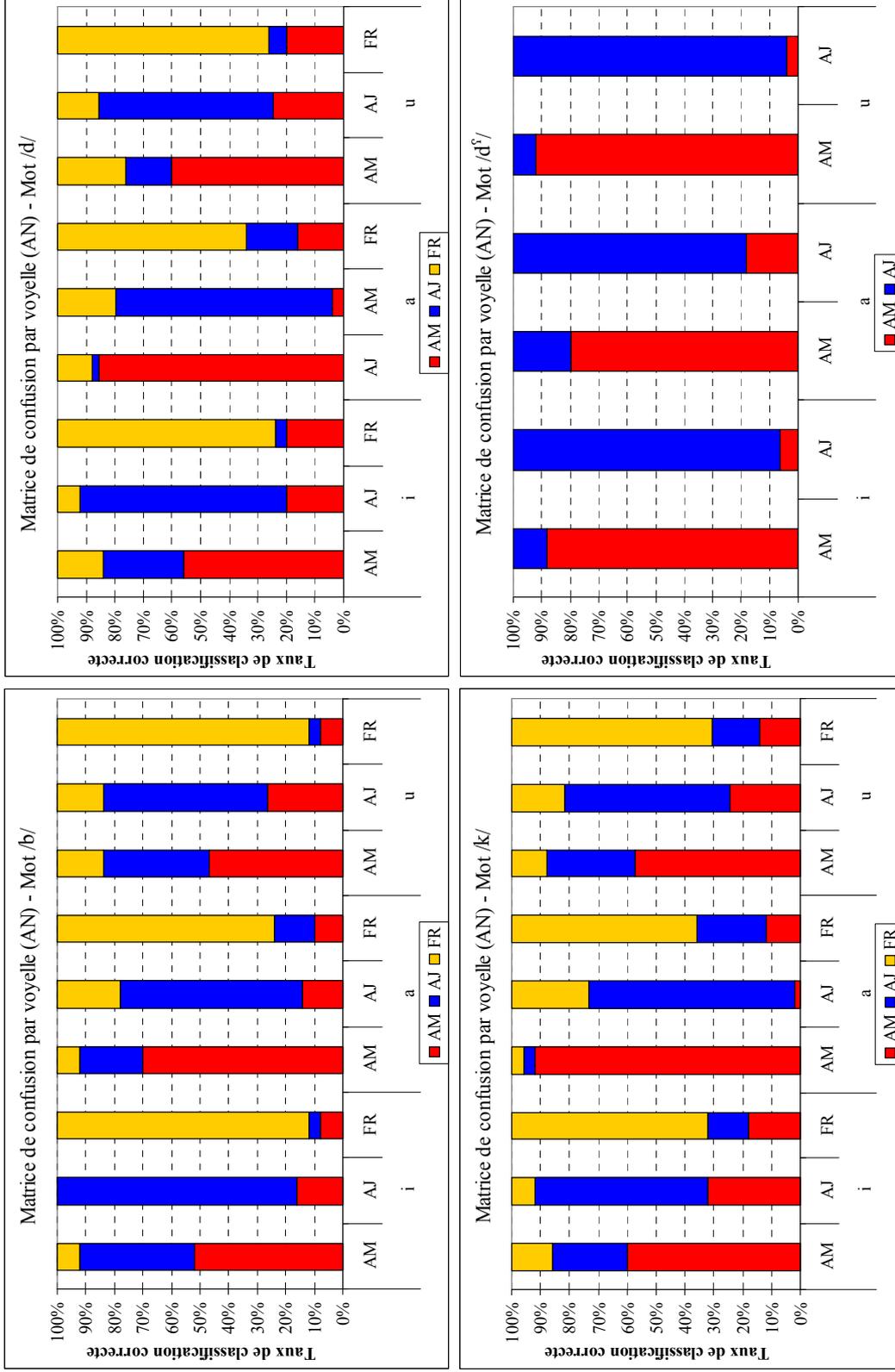


Figure 64 : Matrices de confusion de classification de chaque voyelle /i a u/ en **dynamique** (régression polynomiale 3^{ème} ordre) avec normalisation du temps (AN) dans les trois langues dans la situation de production Mot en fonction du lieu d'articulation des consonnes /b d^f k/ (/d^f/ entre AM & 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 ($F(1, 1140)=0,34$; $p=0,71$).

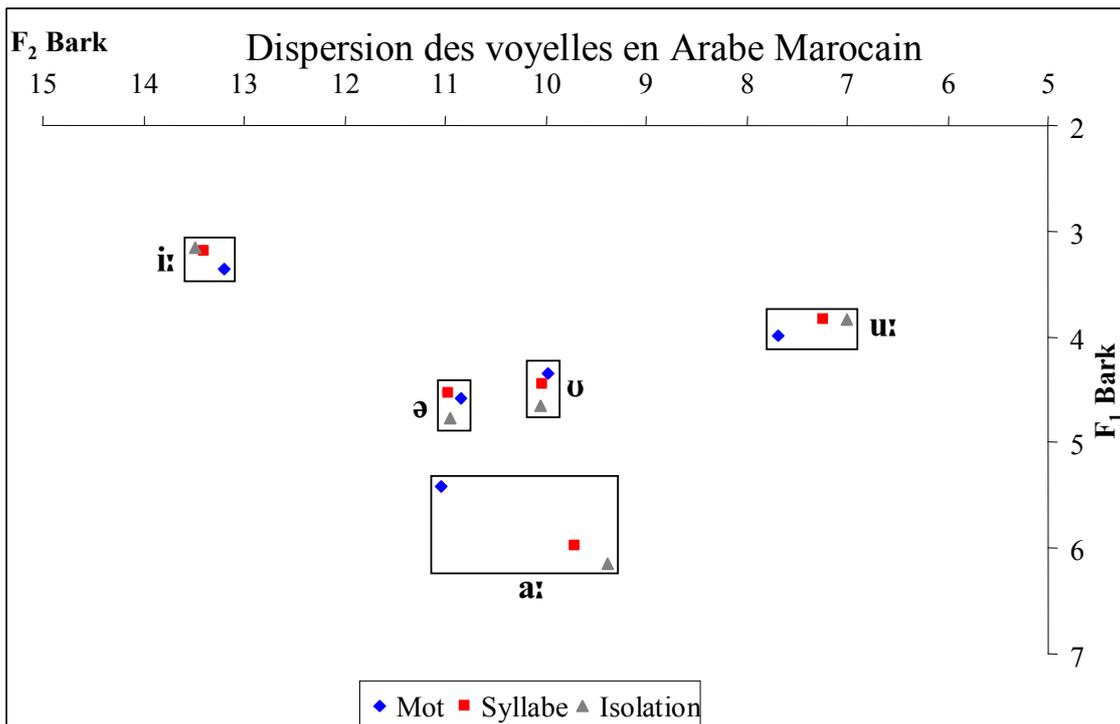


Figure 65 : Espace vocalique en AM en fonction de la situation de production⁶².

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 montre 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 soucis de visibilité pour regrouper les données par timbre vocalique. Leur taille est arbitraire et donc sans lien avec la dispersion des ellipses.

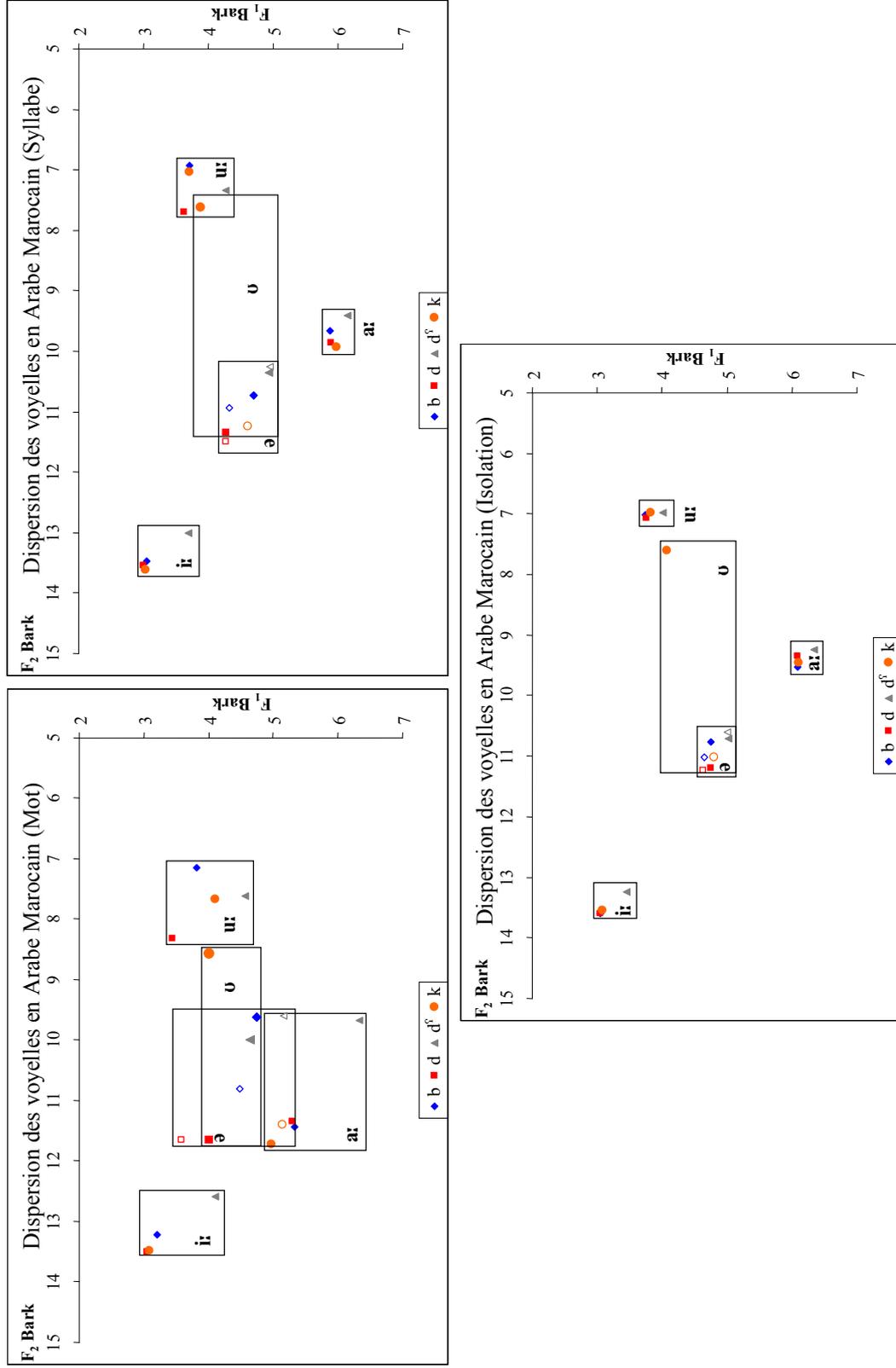


Figure 66 : Espaces vocaliques en AM obtenus en fonction de chaque situation de production et du lieu d'articulation des consonnes /b d d' k/. Les symboles utilisés pour la voyelle /ə/ sont ceux avec un fond vide.

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).

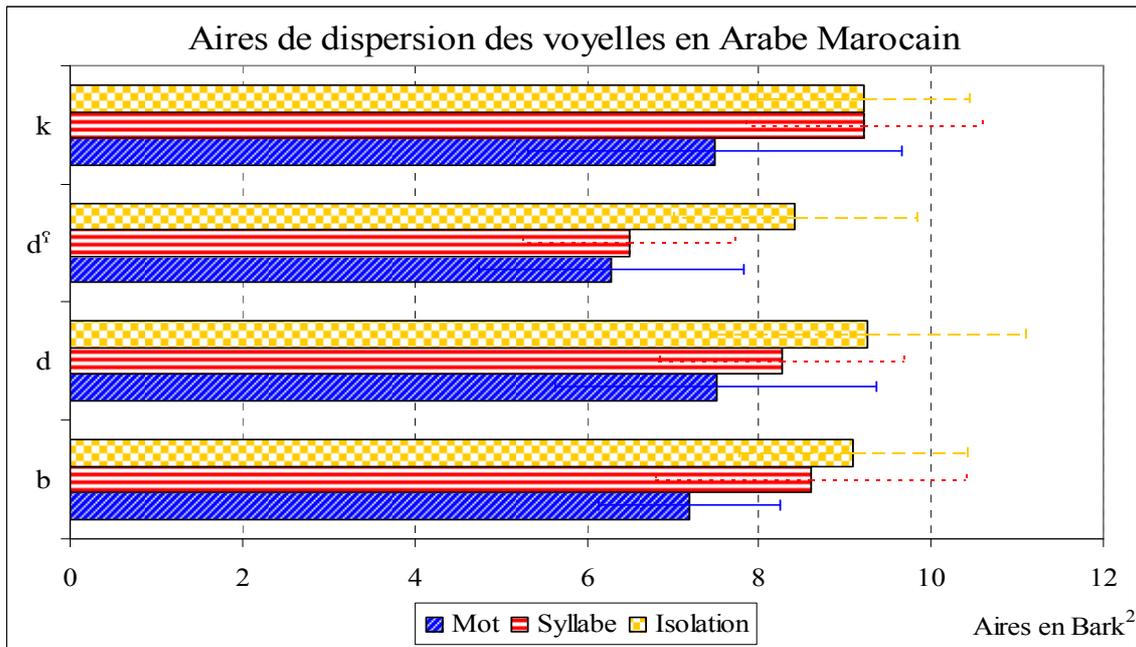


Figure 67 : Aires de dispersion (moyenne et écart-type) des voyelles de l'AM en fonction de la situation de production et du lieu d'articulation.

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.

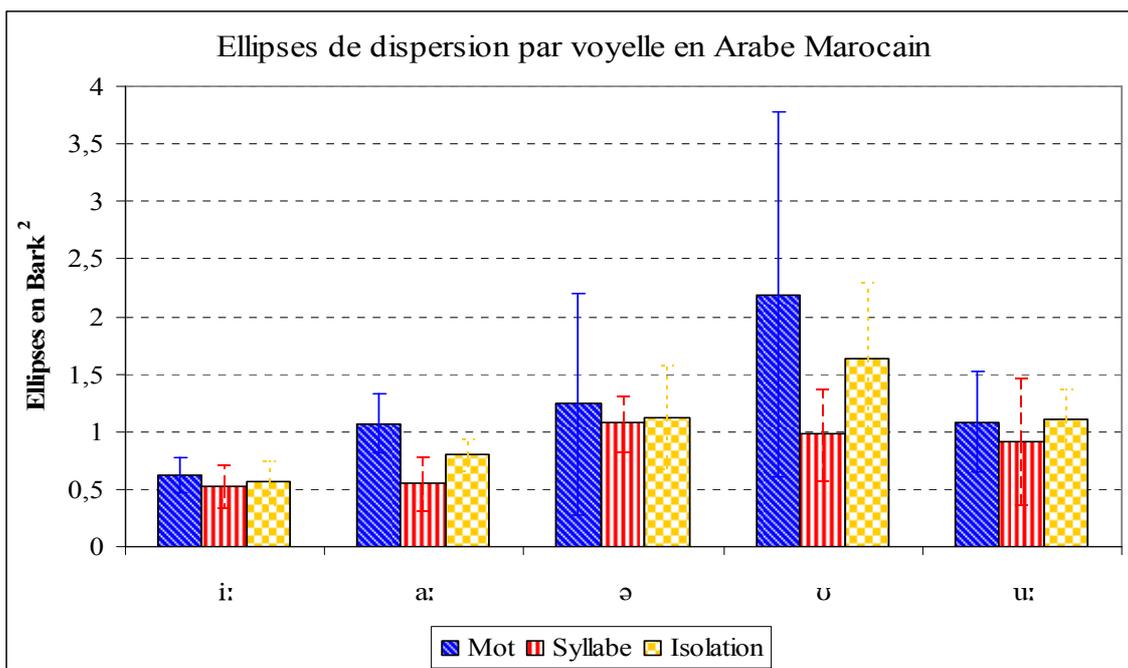


Figure 68 : Ellipses vocaliques en AM en fonction de la situation de production.

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

(Bark ²)		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.

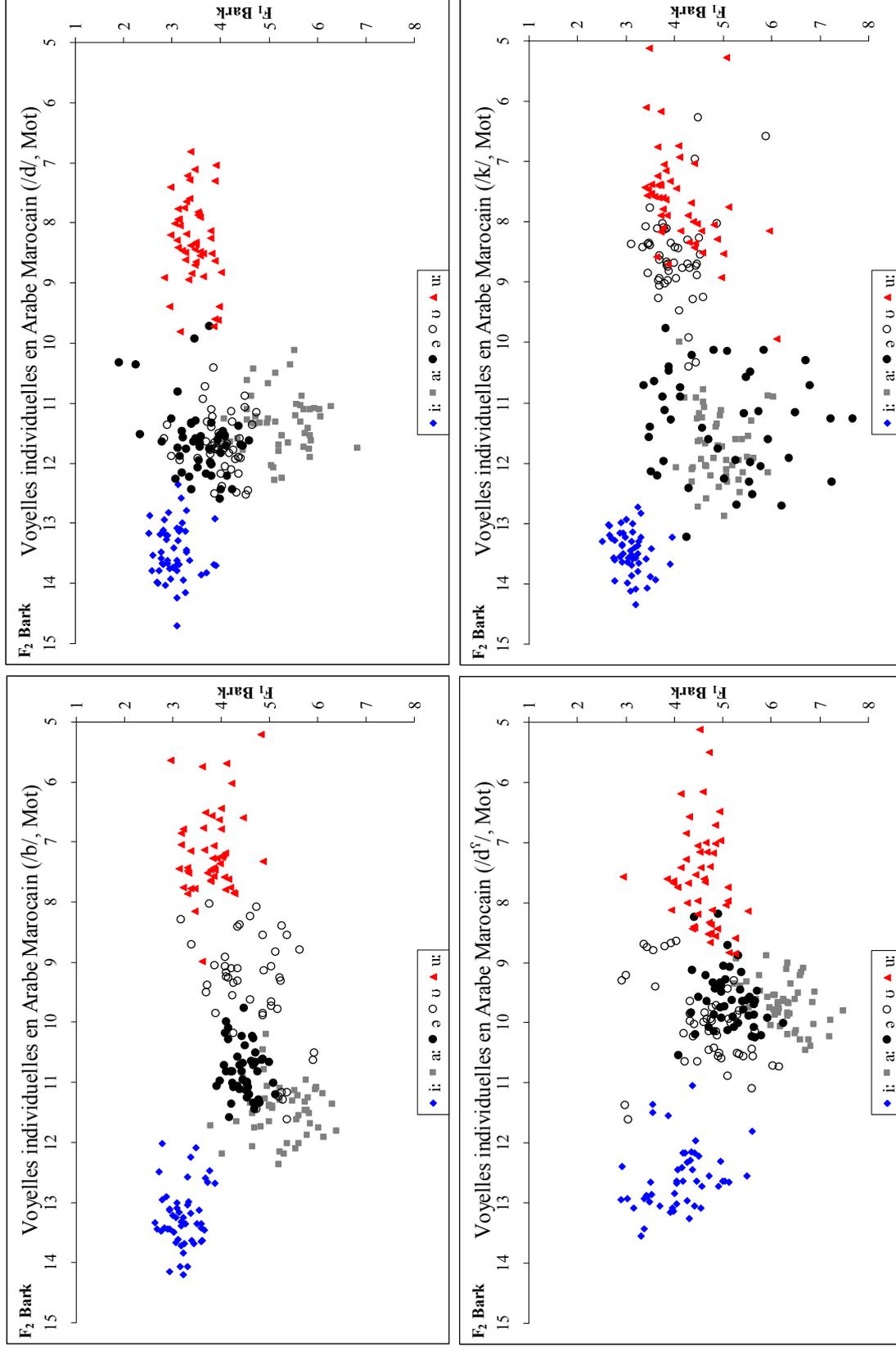


Figure 69 : Dispersion des voyelles en AM à partir des données individuelles, par lieu d'articulation, en Mot.

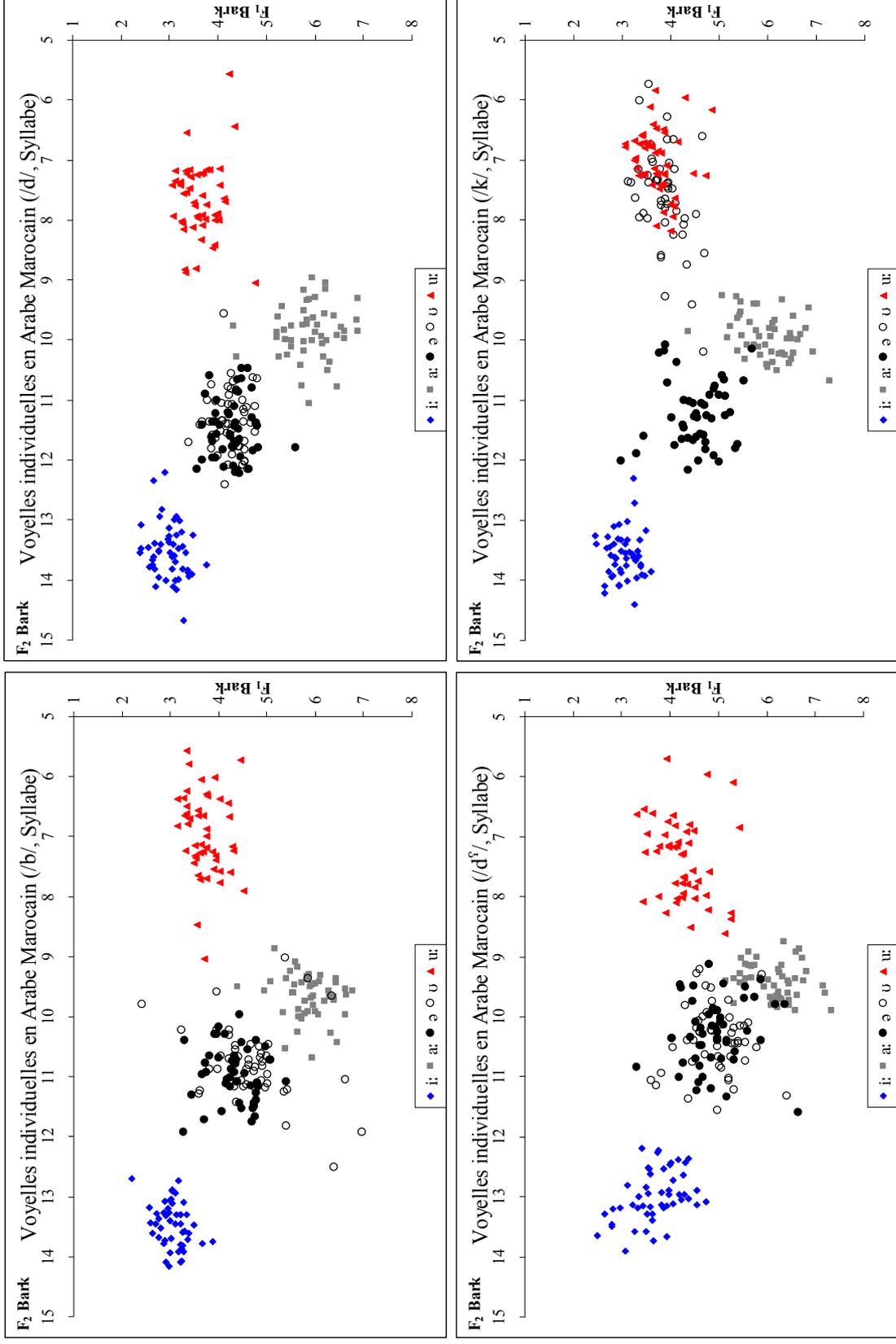


Figure 70 : Dispersion des voyelles en AM à partir des données individuelles, par lieu d'articulation, en Syllabe.

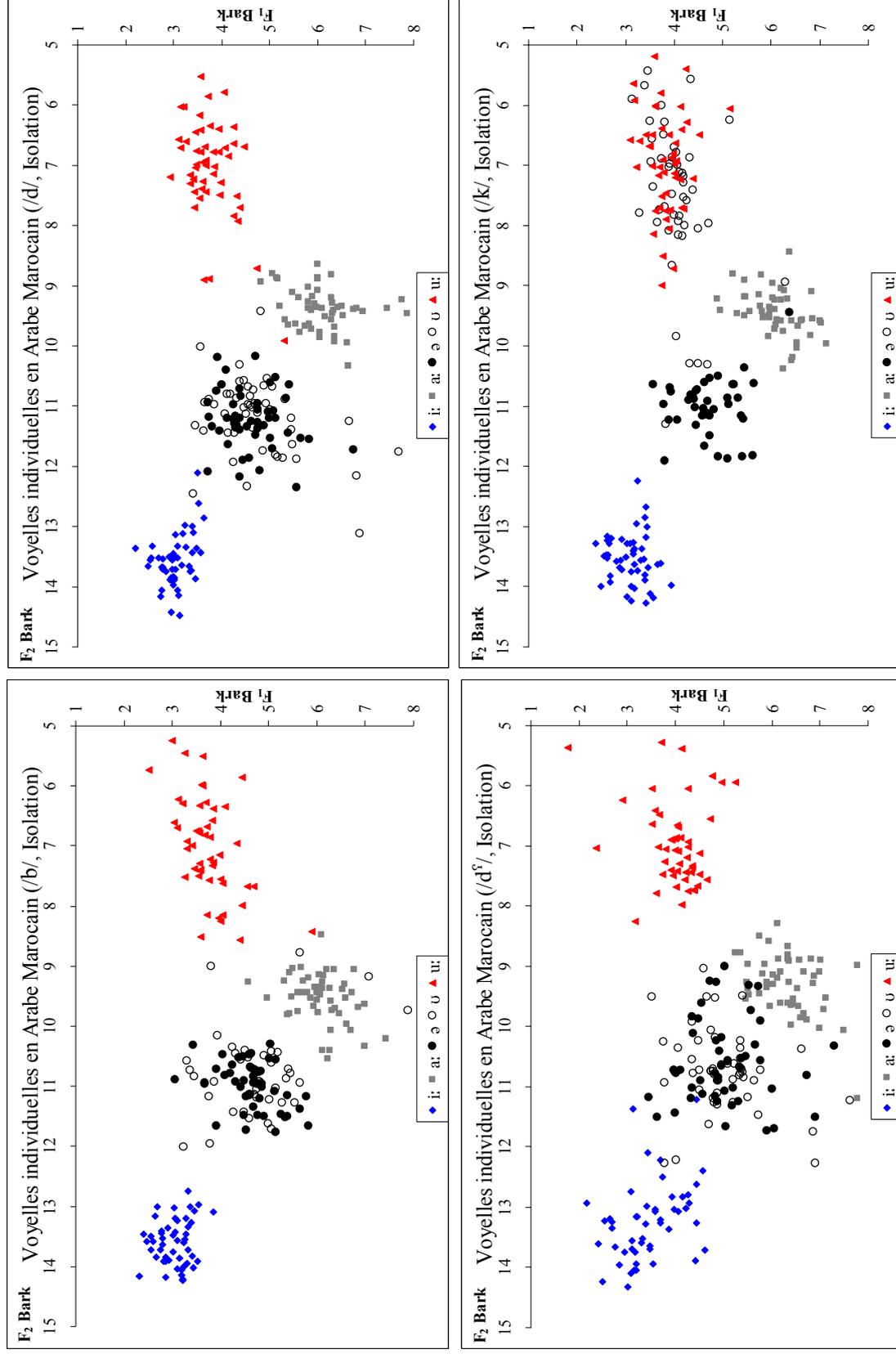


Figure 71 : Dispersion des voyelles en AM à partir des données individuelles, par lieu d'articulation, en Isolation.

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 ($F(3, 338) = 15,08$; $p < 0,001$).

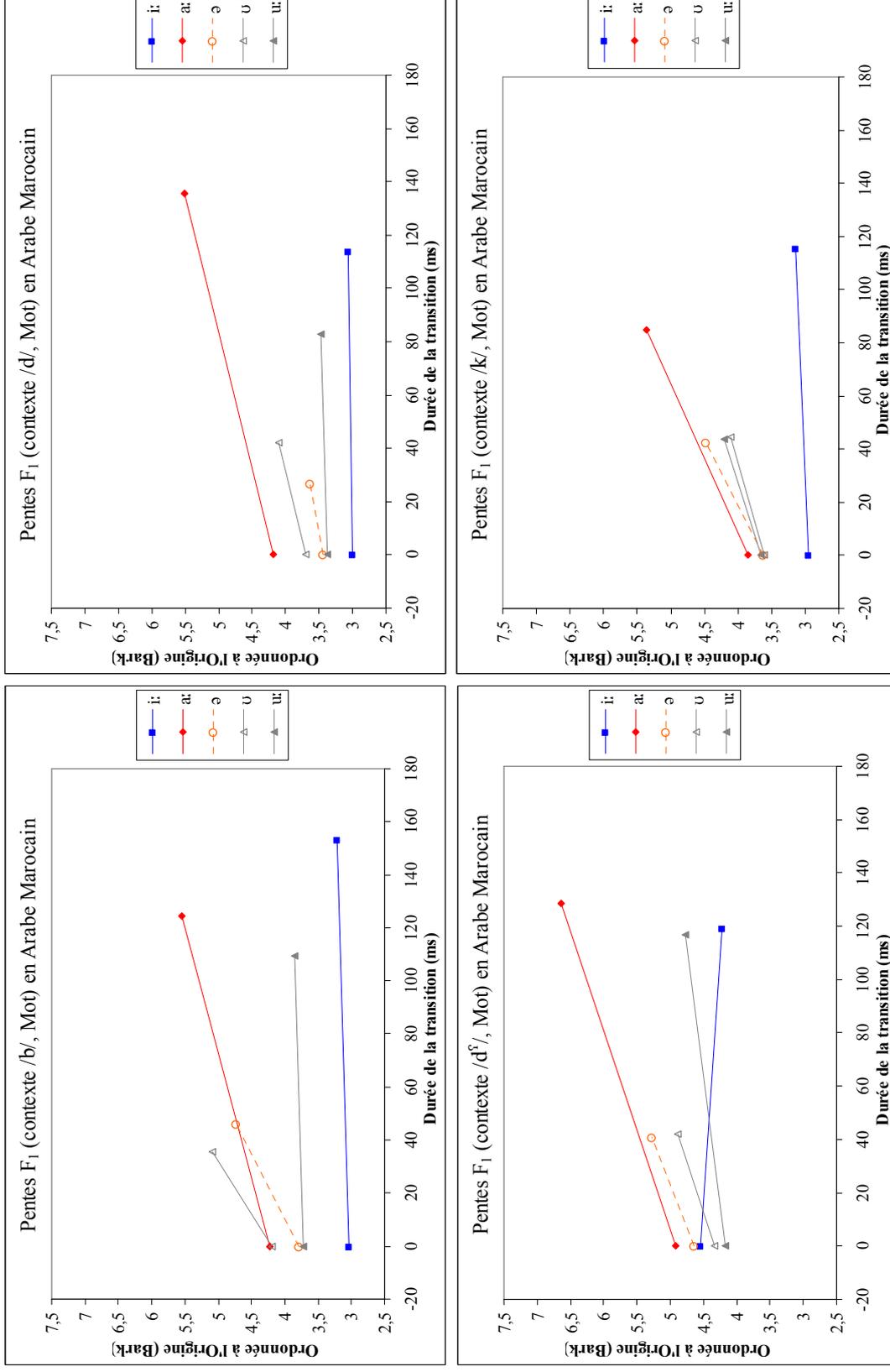


Figure 72 : Pentes formantiques de F₁ des voyelles en AM sans normalisation du temps, en fonction du lieu d'articulation en Mot.

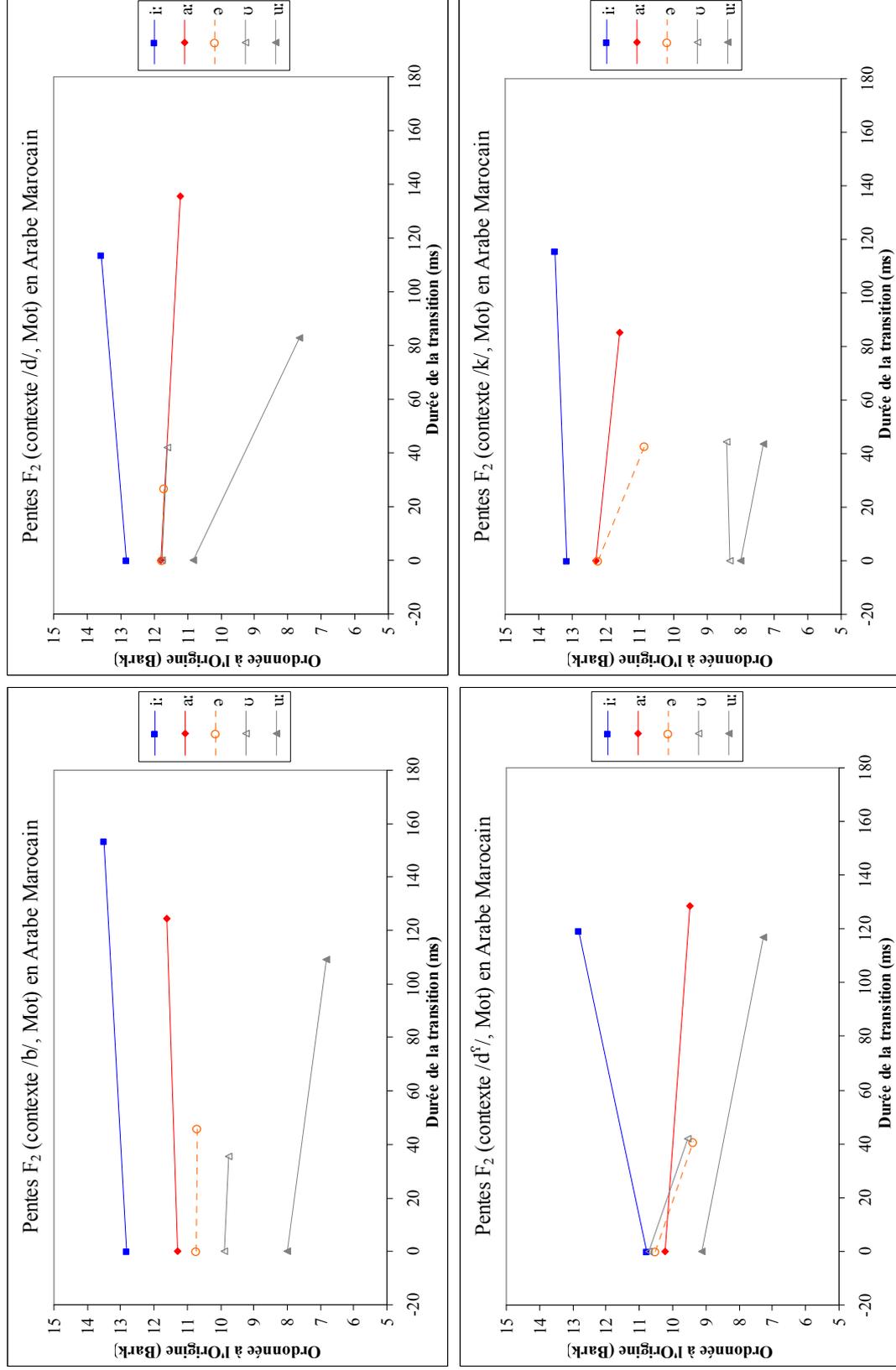


Figure 73 : Pentes formantiques de F₂ des voyelles en AM sans normalisation du temps, en fonction du lieu d'articulation en Mot.

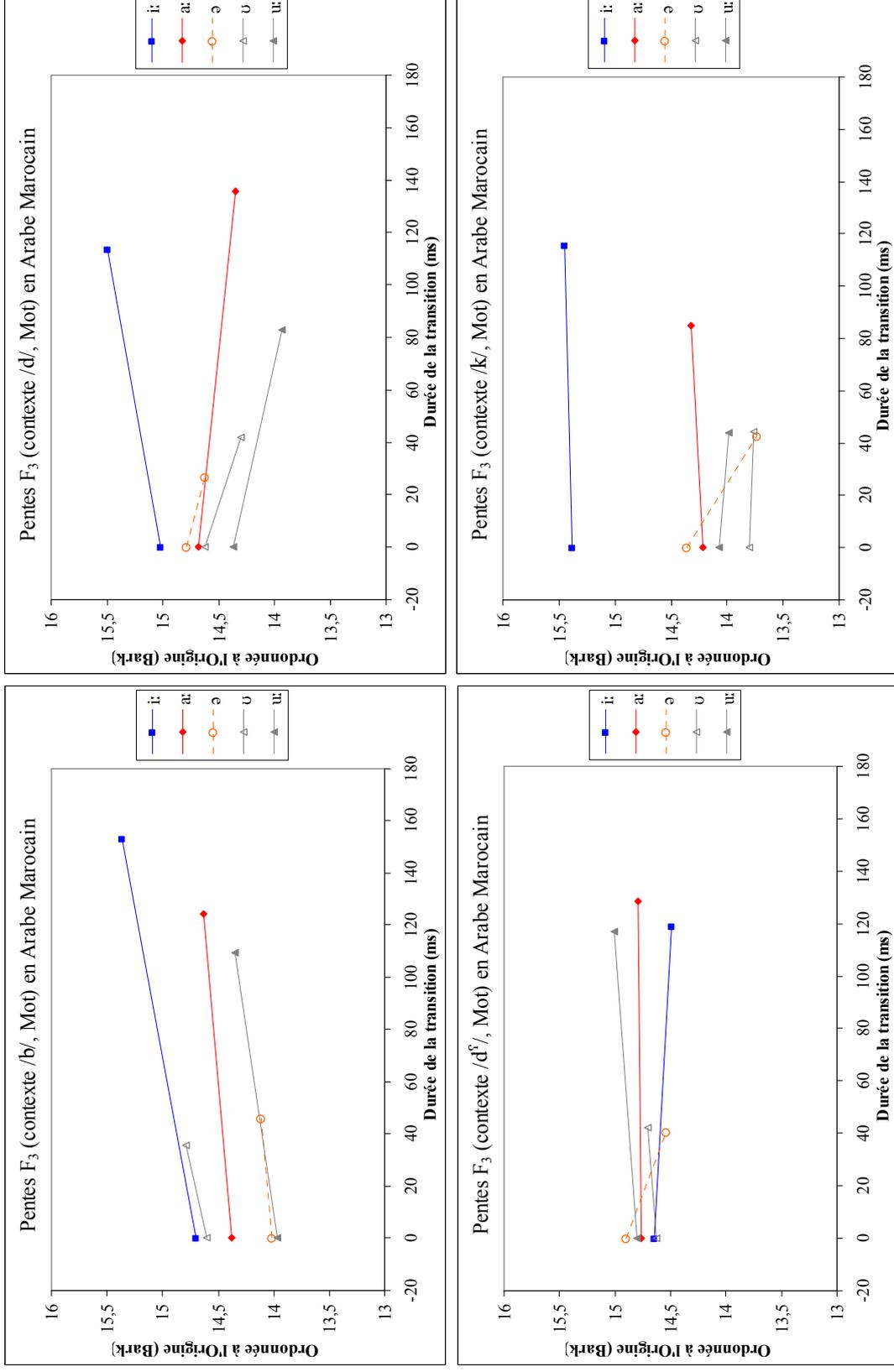


Figure 74 : Pentas formantiques de F₃ des voyelles en AM sans normalisation du temps, en fonction du lieu d'articulation en Mot.

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^ɪ/, 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 ^ɪ /	/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.

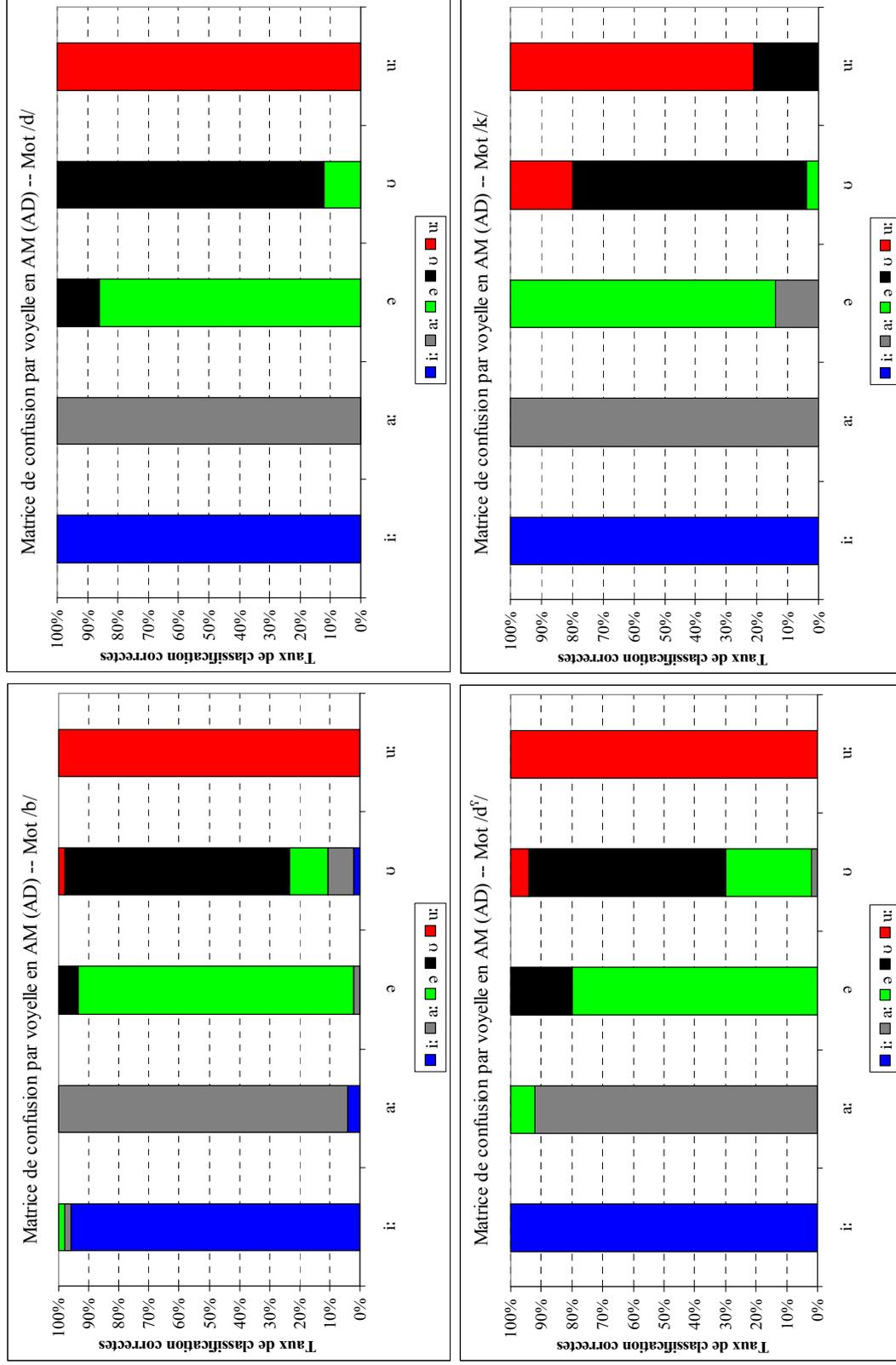


Figure 75 : Matrices de confusion de classification des voyelles de l'AM en **statique** avec la durée vocalique (AD) en situation de production Mot et en fonction du lieu d'articulation.

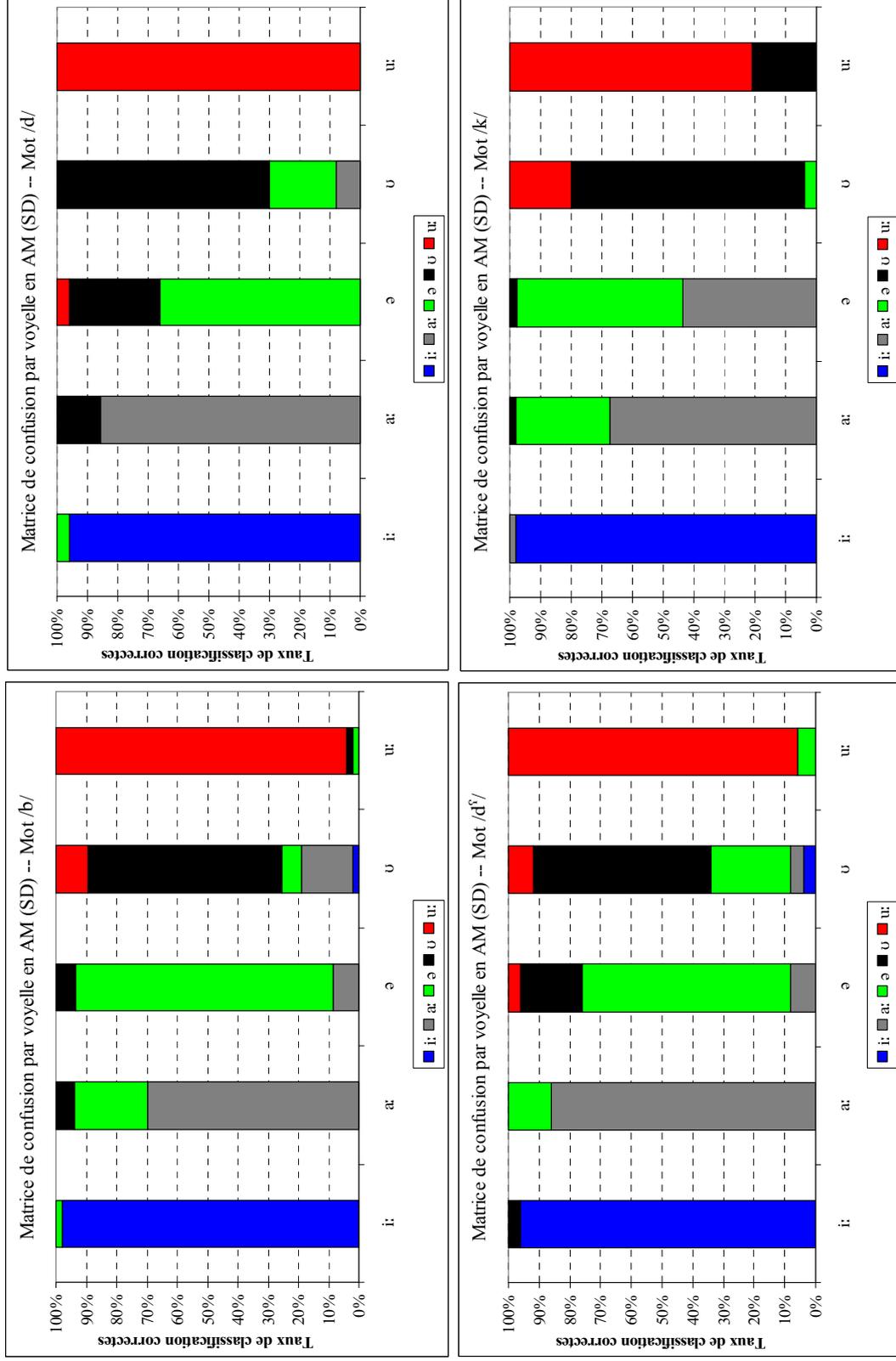


Figure 76 : Matrices de confusion de classification des voyelles de l'AM en **statique** sans la durée vocalique (SD) en situation de production Mot et en fonction du lieu d'articulation.

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.

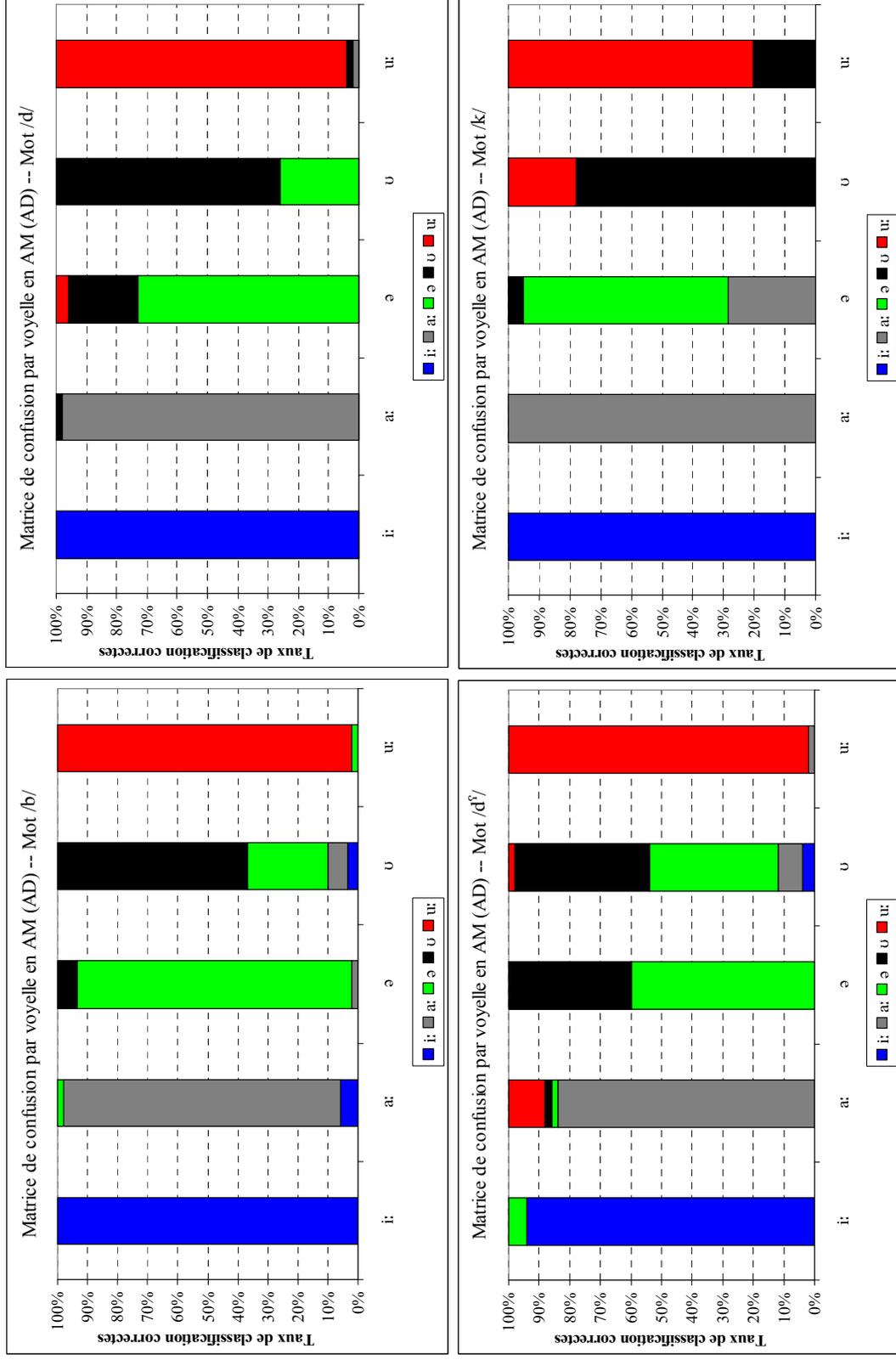


Figure 77 : Taux de classification des voyelles de l'AM en dynamique (régression linéaire) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

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).

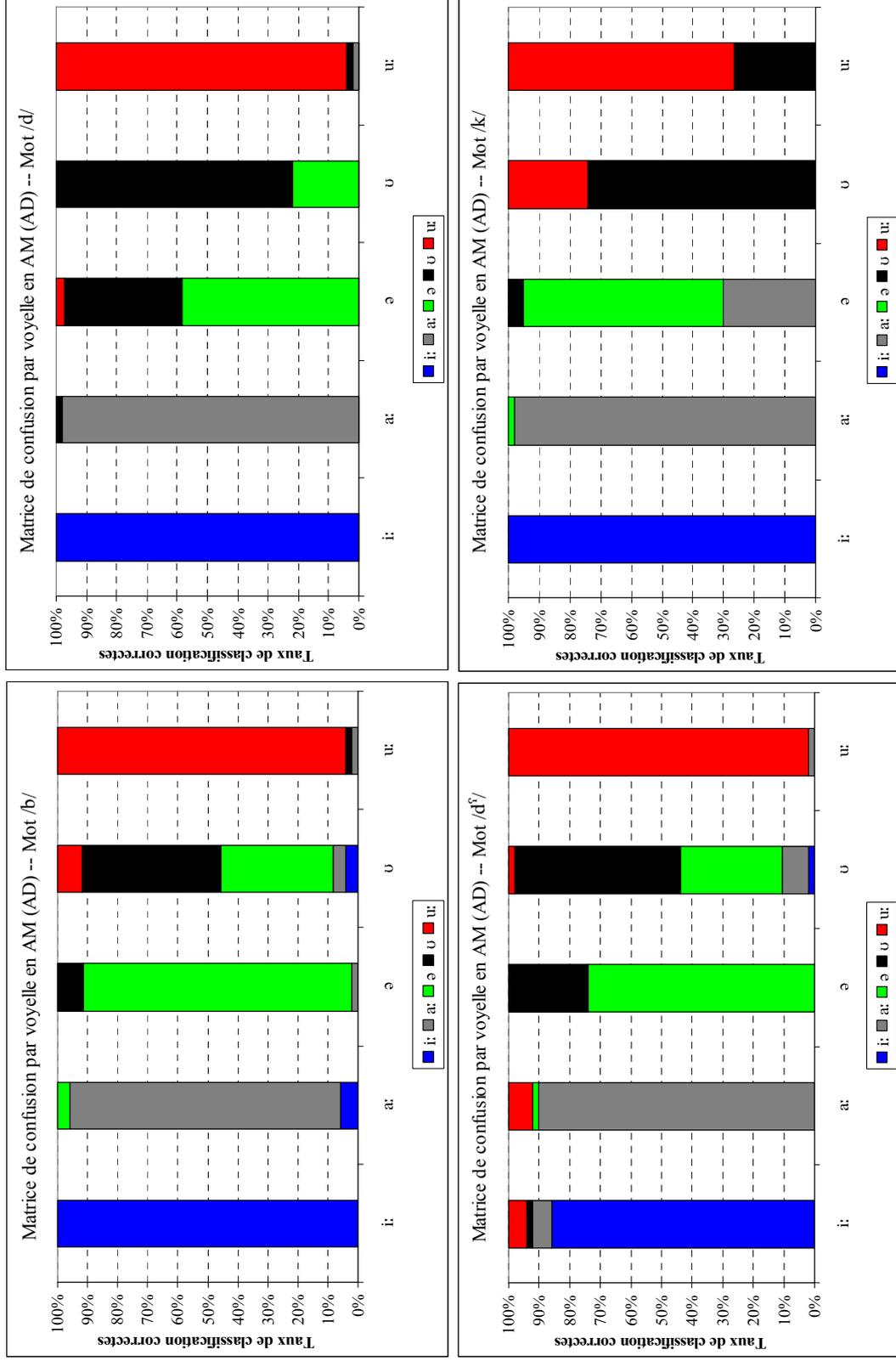


Figure 78 : Matrices de confusion de classification des voyelles de l'AM en **dynamique** (régression polynomiale de 2^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

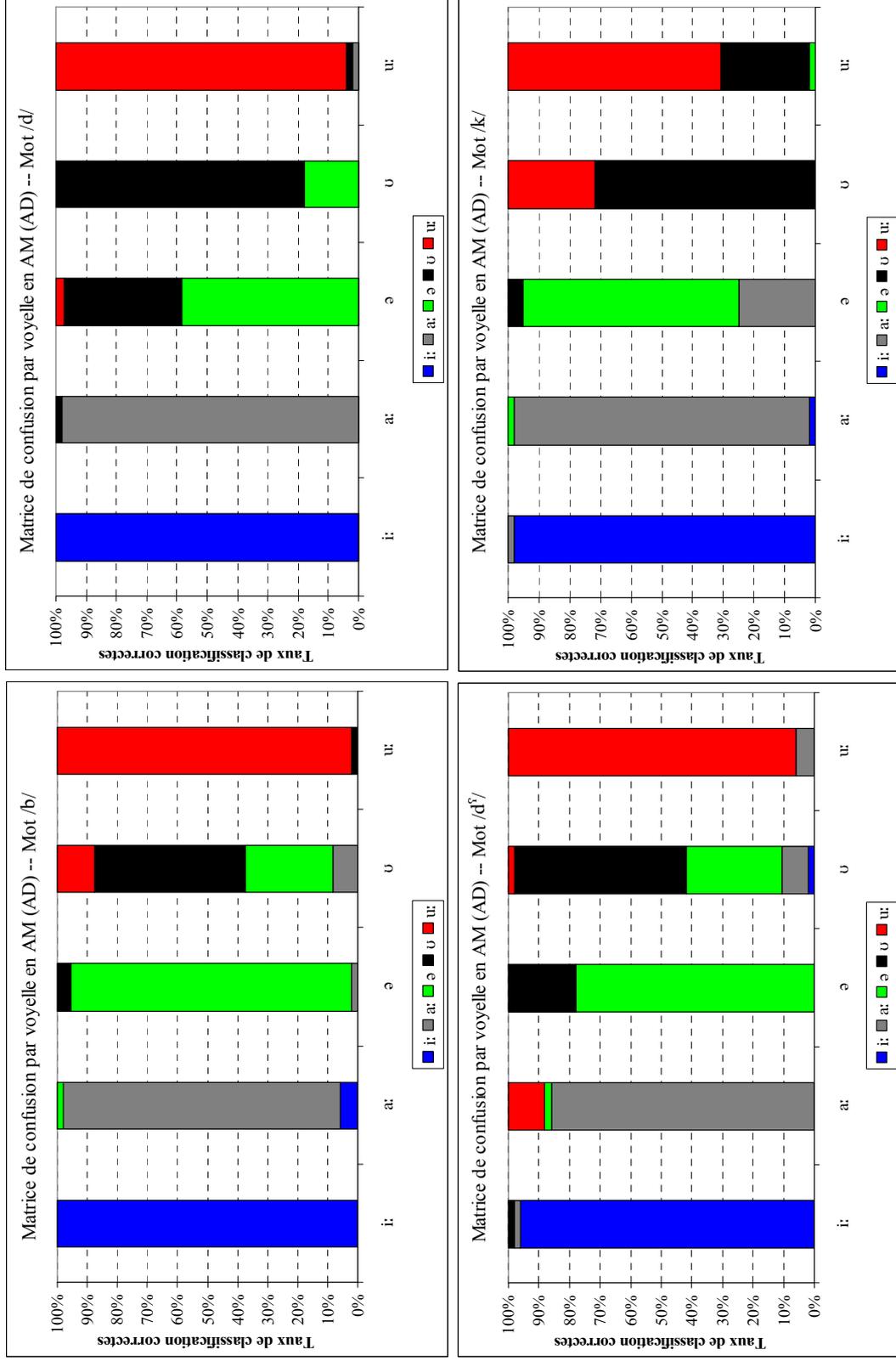


Figure 79 : Matrices de confusion de classification des voyelles de l'AM en **dynamique** (régression polynomiale de 3^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

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 ; $p < 0,05$).

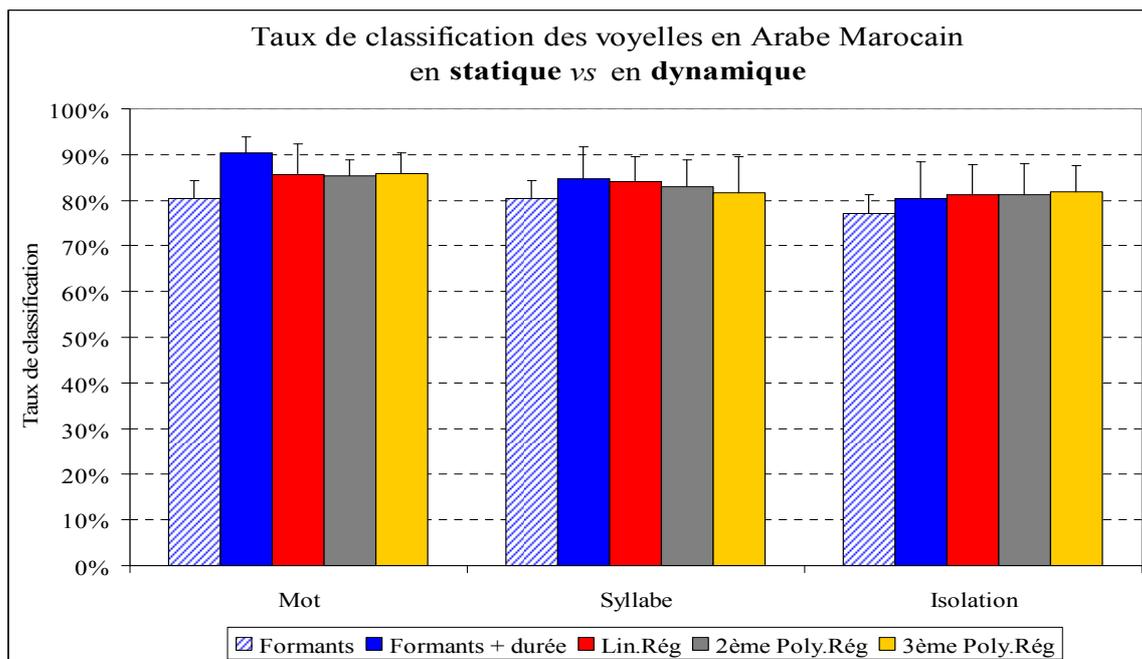


Figure 80 : Taux de classification correcte moyens des voyelles de l'AM produites dans les contextes /b d d^f k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur). Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} 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, ($F(14, 4190) = 7,86$; $p < 0,001$).

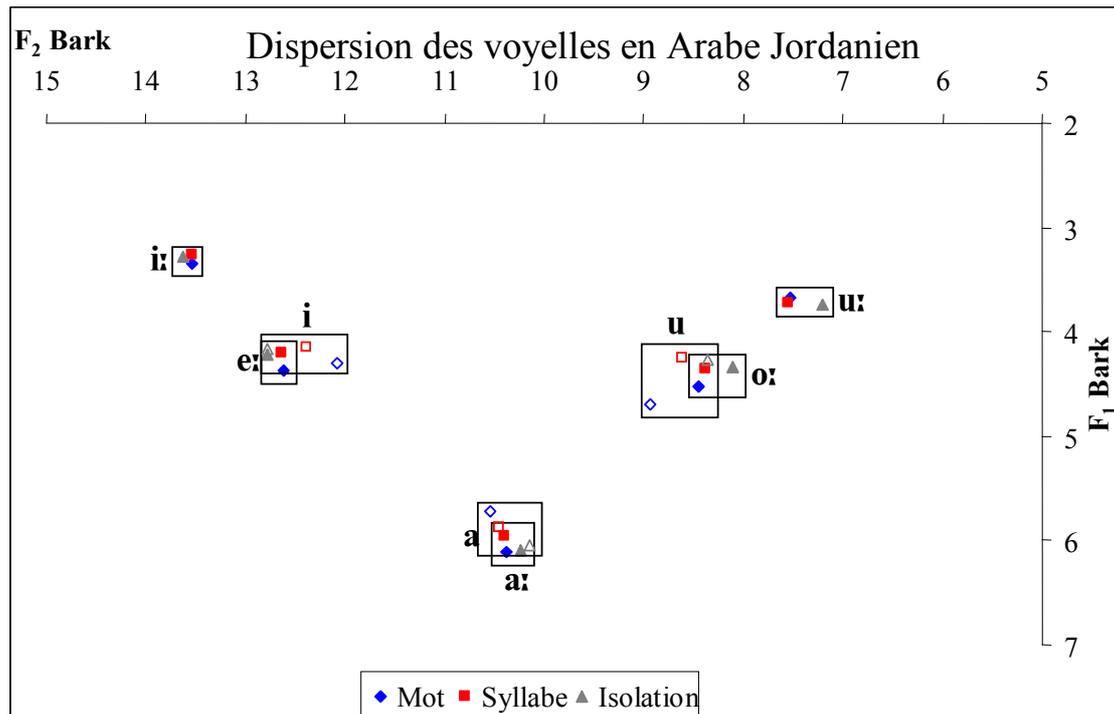


Figure 81 : Espace vocalique en AJ en fonction de la situation de production. Les voyelles brèves sont représentées par les symboles avec les fonds vides.

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 ($F(6, 4190) = 10,68$; $p < 0,001$).

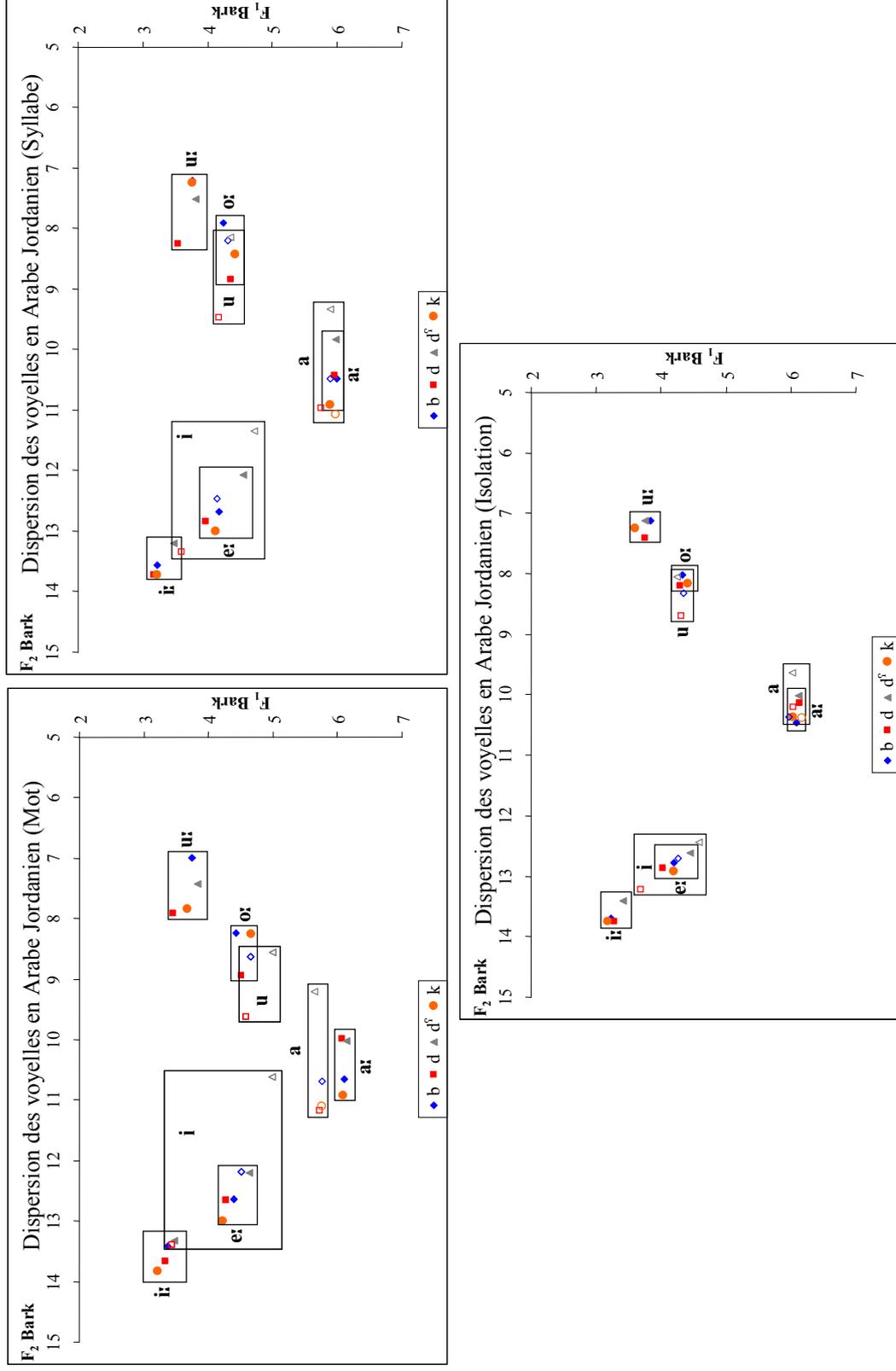


Figure 82 : Espaces vocaliques en AJ en fonction de chaque situation de production et du lieu d'articulation des consonnes /b d d' k/. Les voyelles brèves sont représentées par les symboles avec les fonds vides.

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 ($F(6, 108)=0,32 ; p=0,93$)).

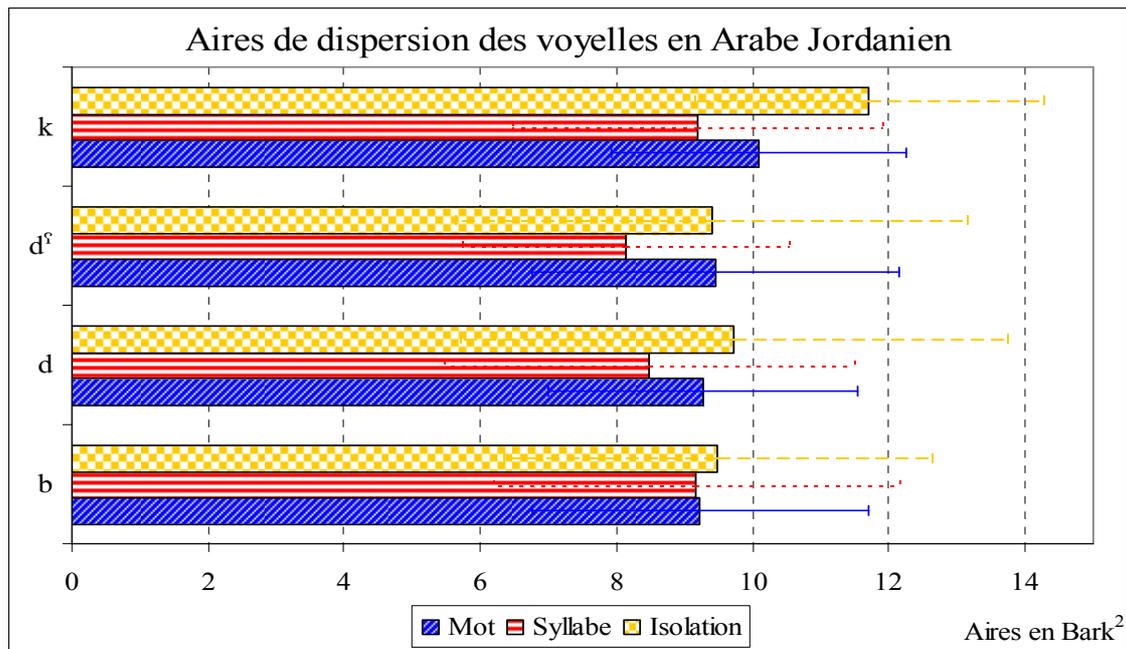


Figure 83 : Aires de dispersion (moyenne et écart-type) des voyelles de l'AJ en fonction de la situation de production et du lieu d'articulation.

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:/).

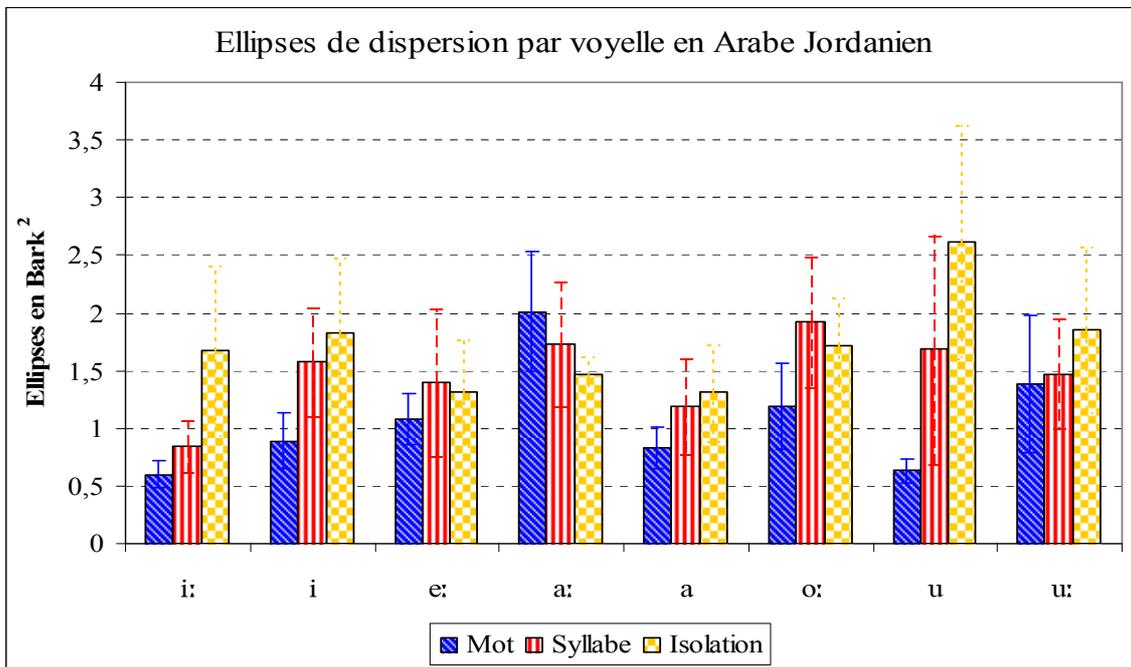


Figure 84 : Ellipses de dispersion (moyenne et écart-type) des voyelles en fonction de la situation de production en AJ.

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.

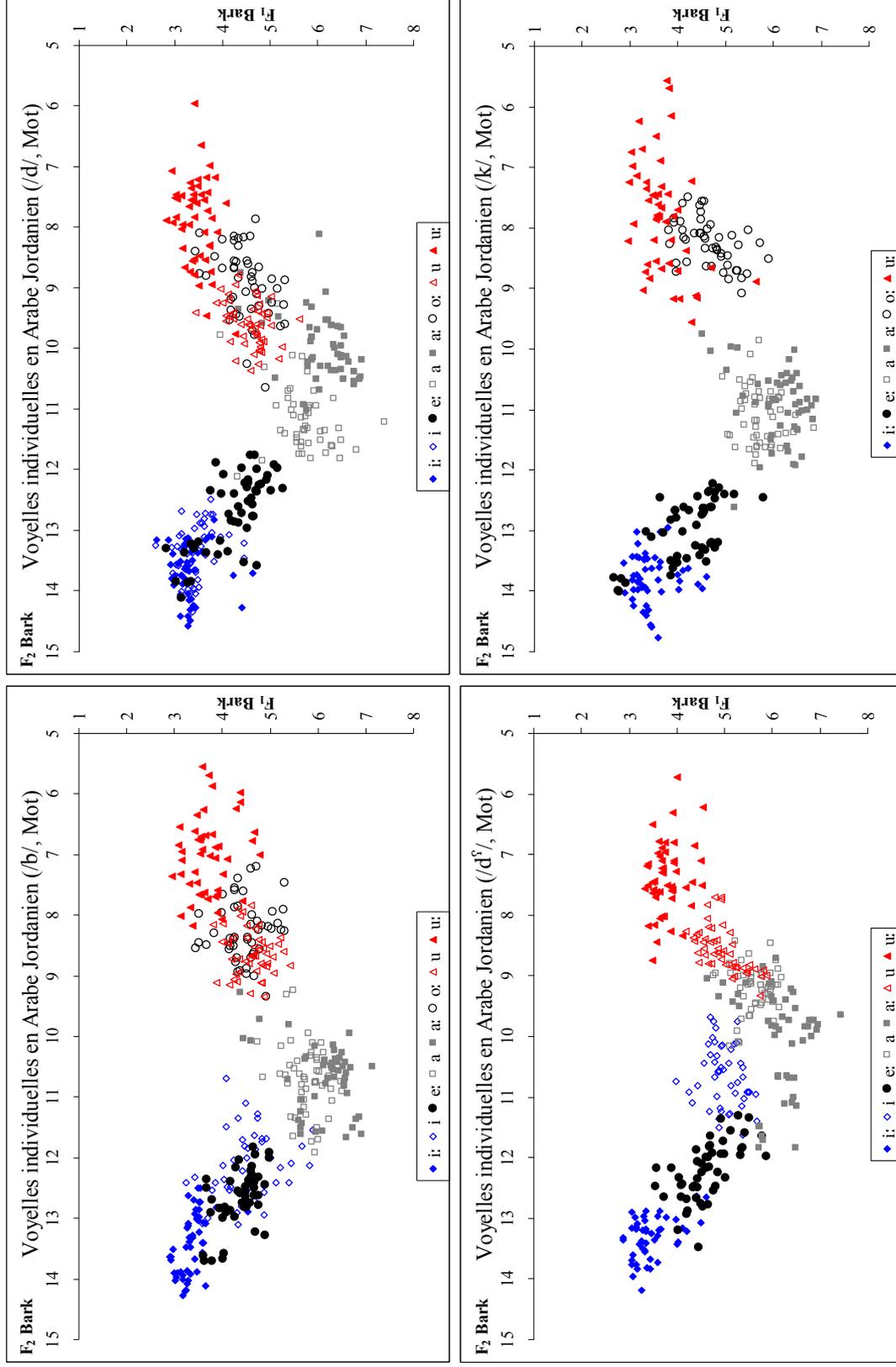


Figure 85 : Dispersion des voyelles en AJ à partir des données individuelles, par lieu d'articulation, en Mot.

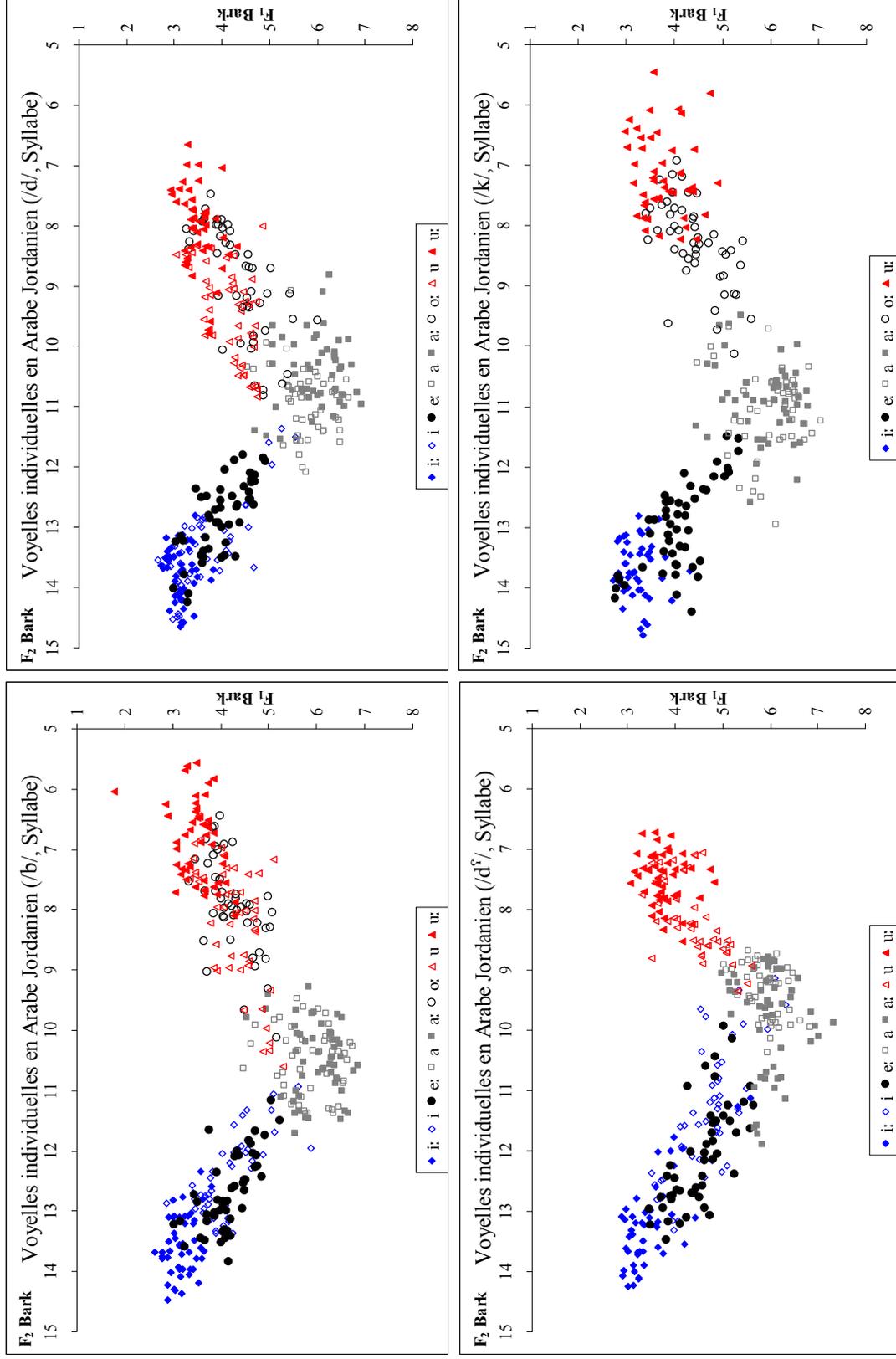


Figure 86 : Dispersion des voyelles en AJ à partir des données individuelles, par lieu d'articulation, en Syllabe.

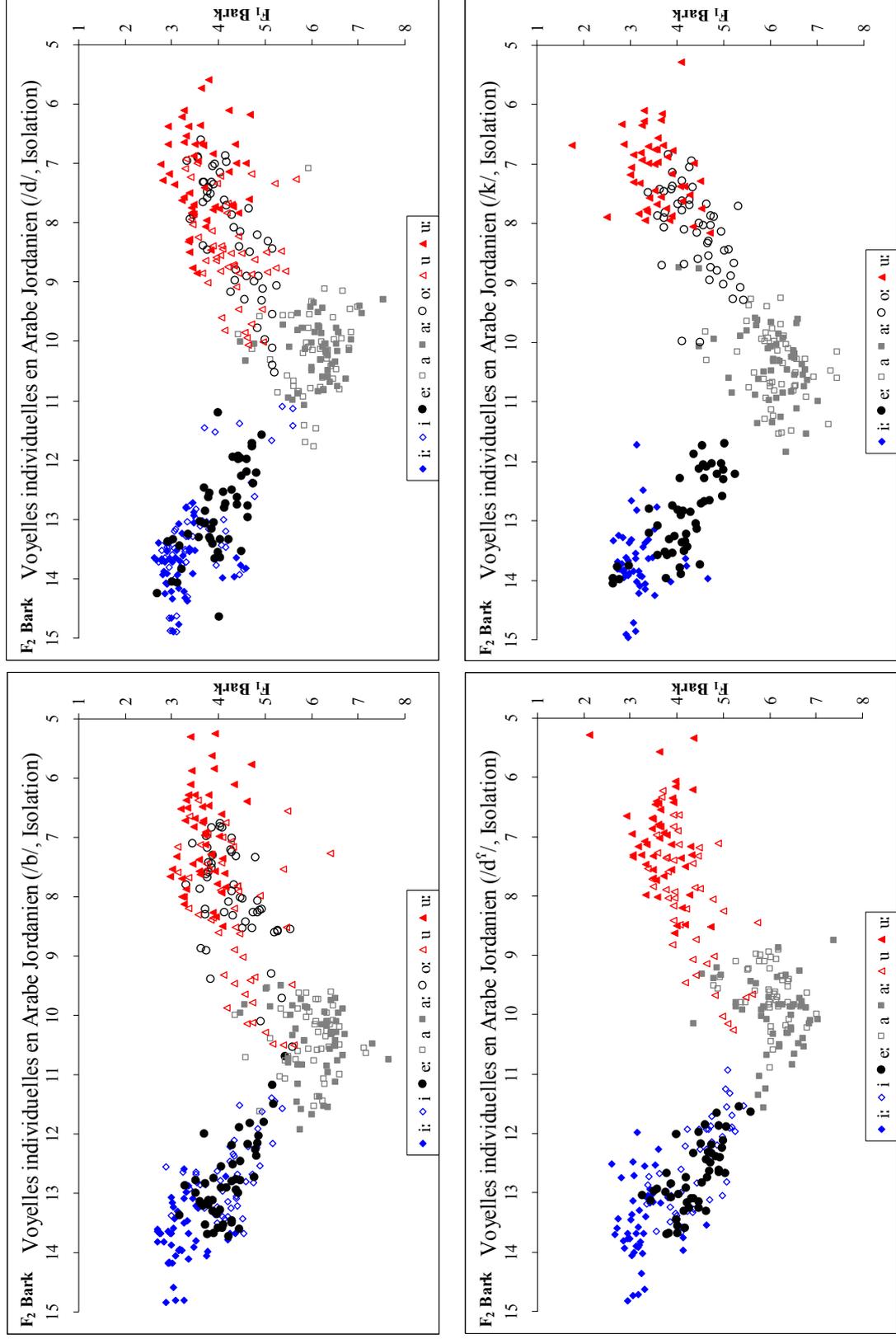


Figure 87 : Dispersion des voyelles en AJ à partir des données individuelles, par lieu d'articulation, en Isolation.

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.

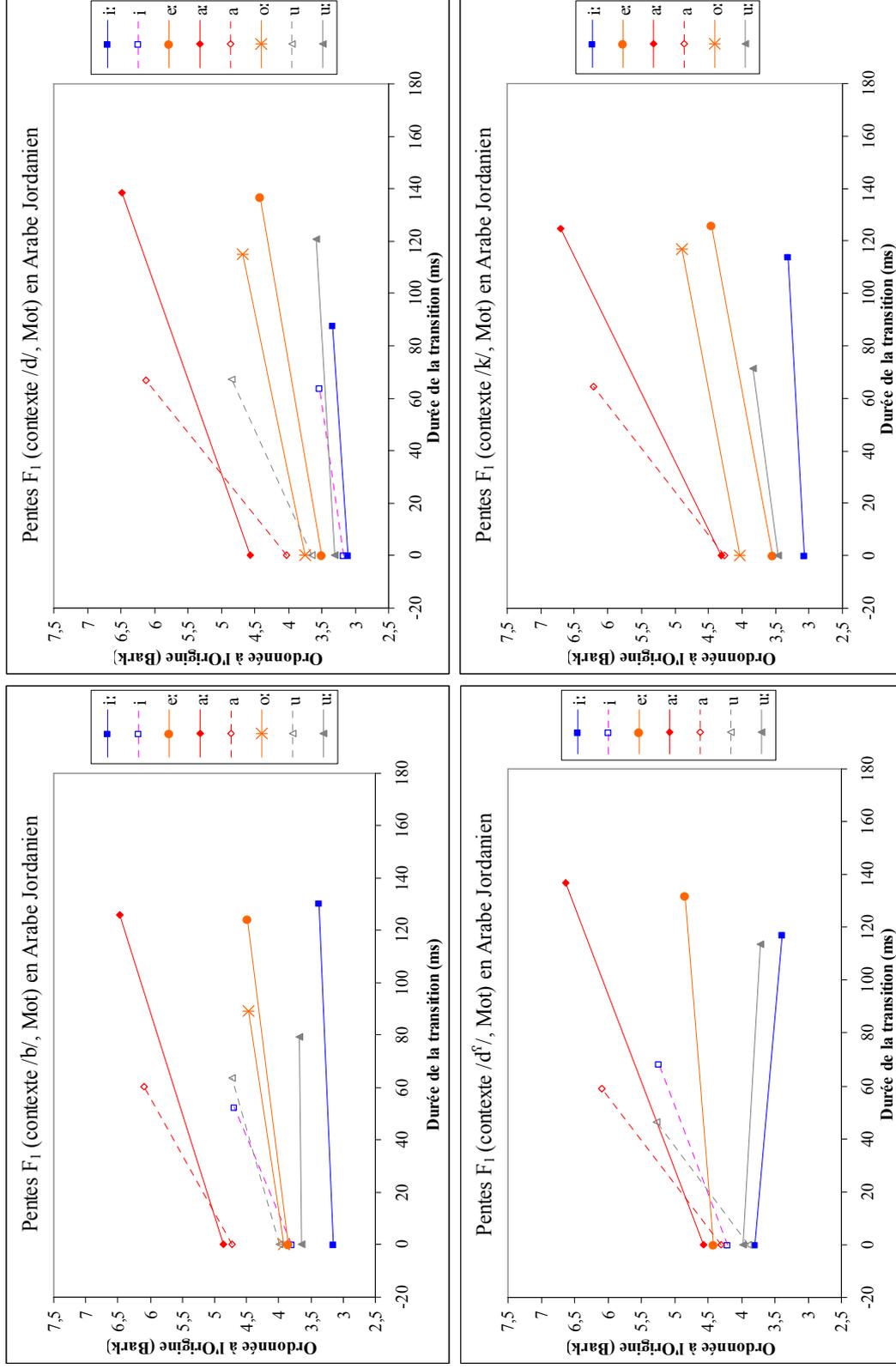


Figure 88 : Pentas formantiques de F₁ des voyelles en AJ sans normalisation du temps, en fonction du lieu d'articulation en Mot.

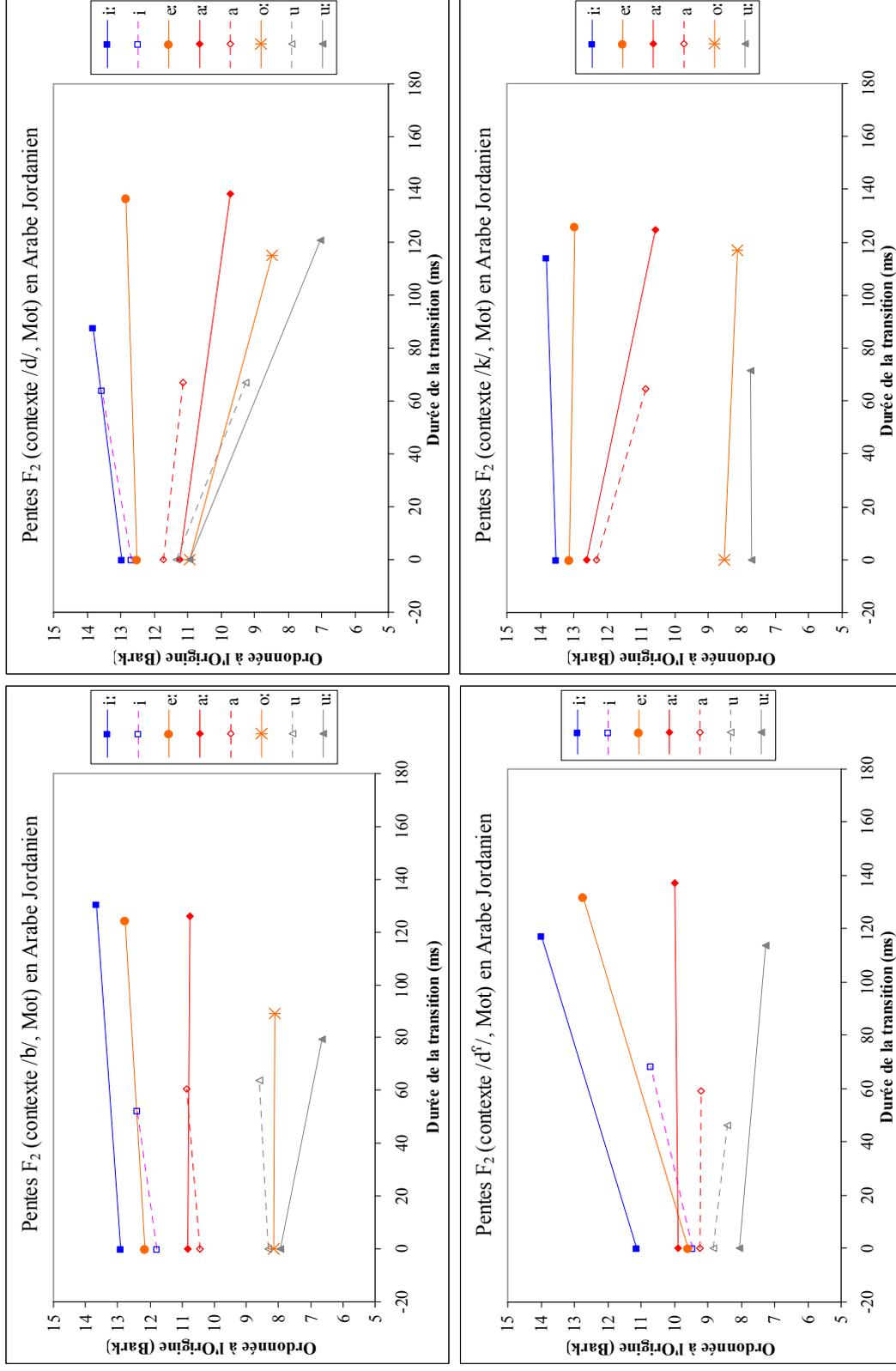


Figure 89 : Pentas formantiques de F₂ des voyelles en AJ sans normalisation du temps, en fonction du lieu d'articulation en Mot.

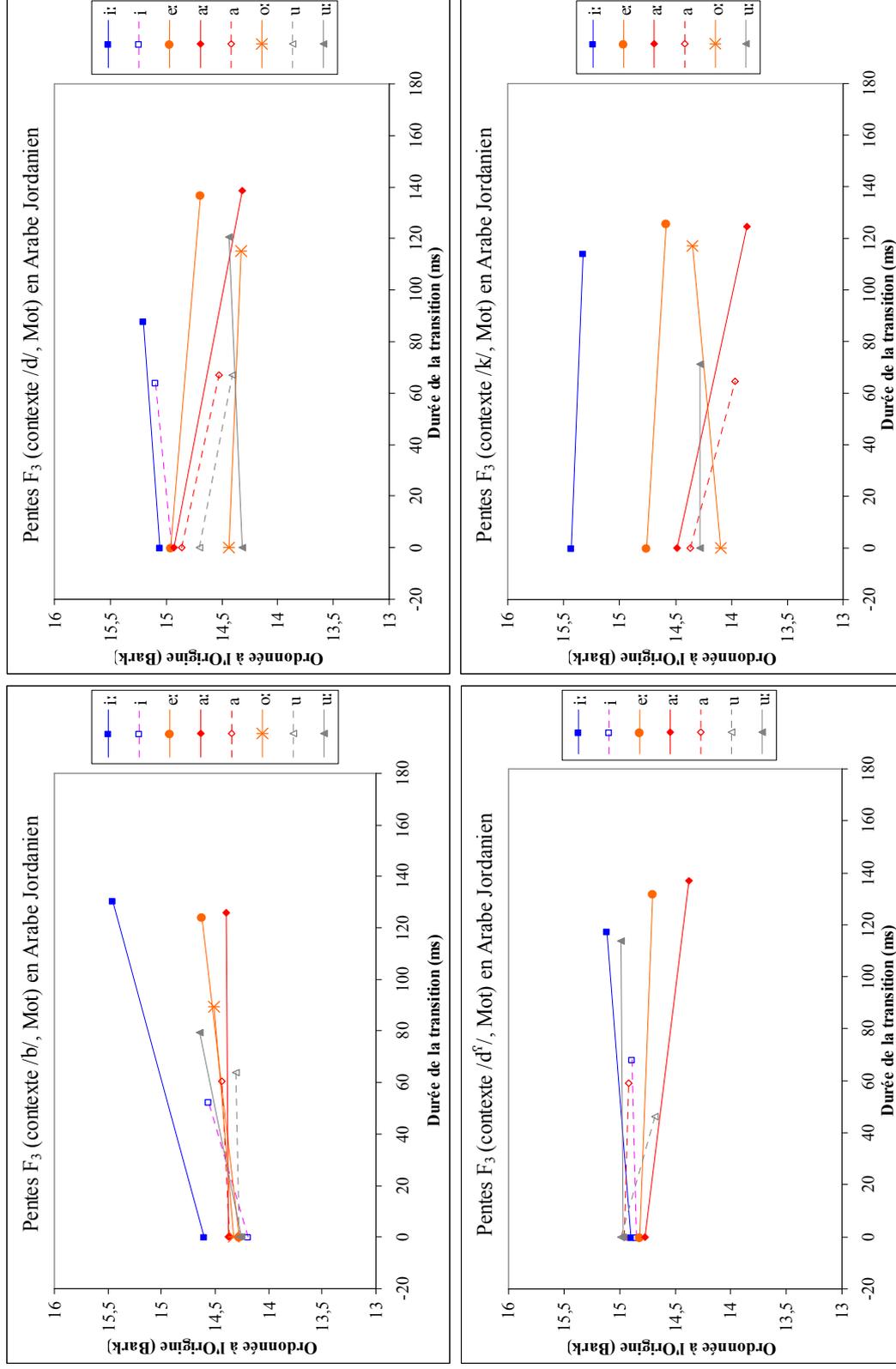


Figure 90 : Pentas formantiques de F₃ des voyelles en AJ sans normalisation du temps, en fonction du lieu d'articulation en Mot.

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).

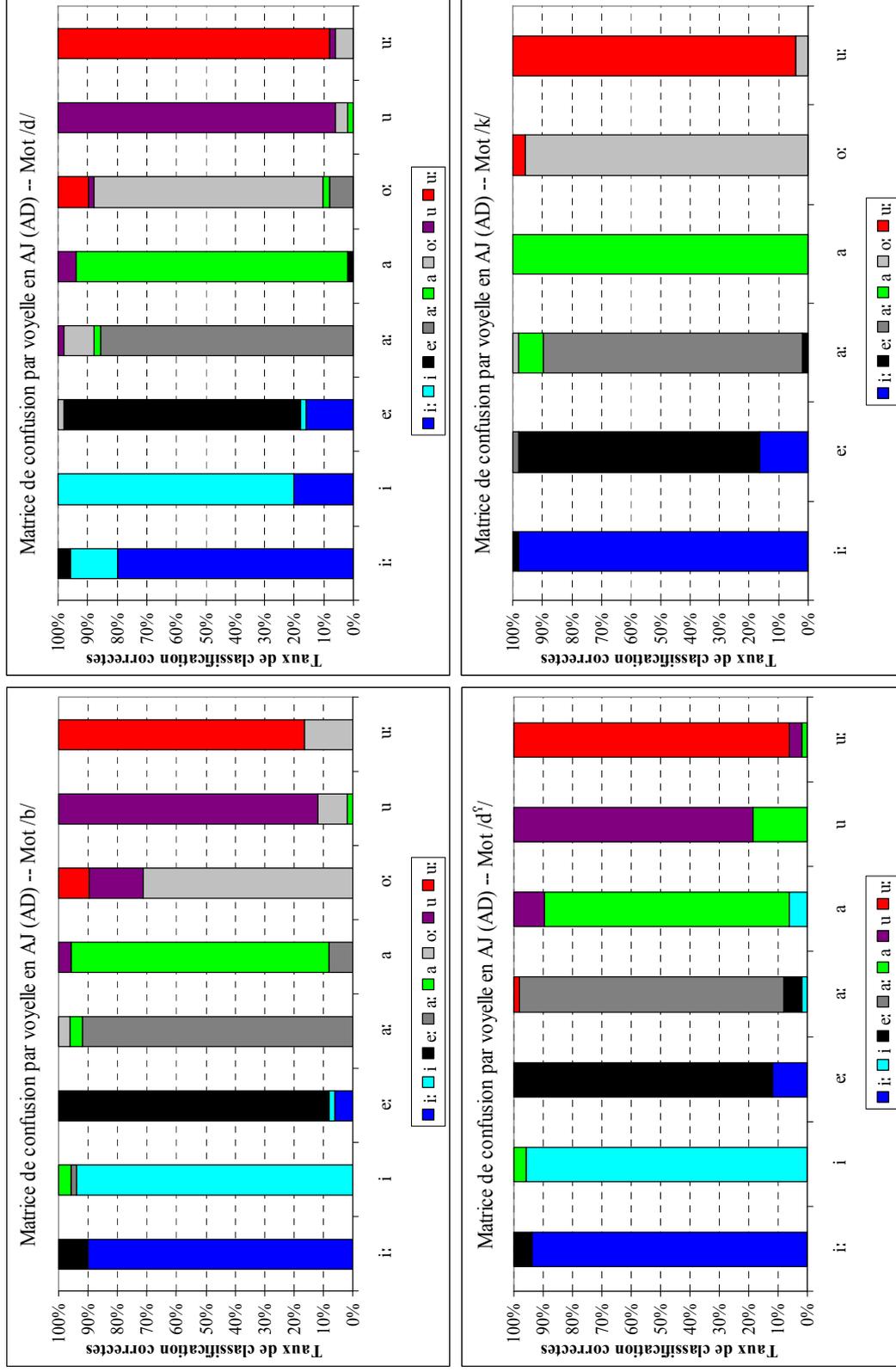


Figure 91 : Matrices de confusion de classification des voyelles de l'AJ en **statique** avec la durée vocalique (AD) en situation de production Mot et en fonction du lieu d'articulation.

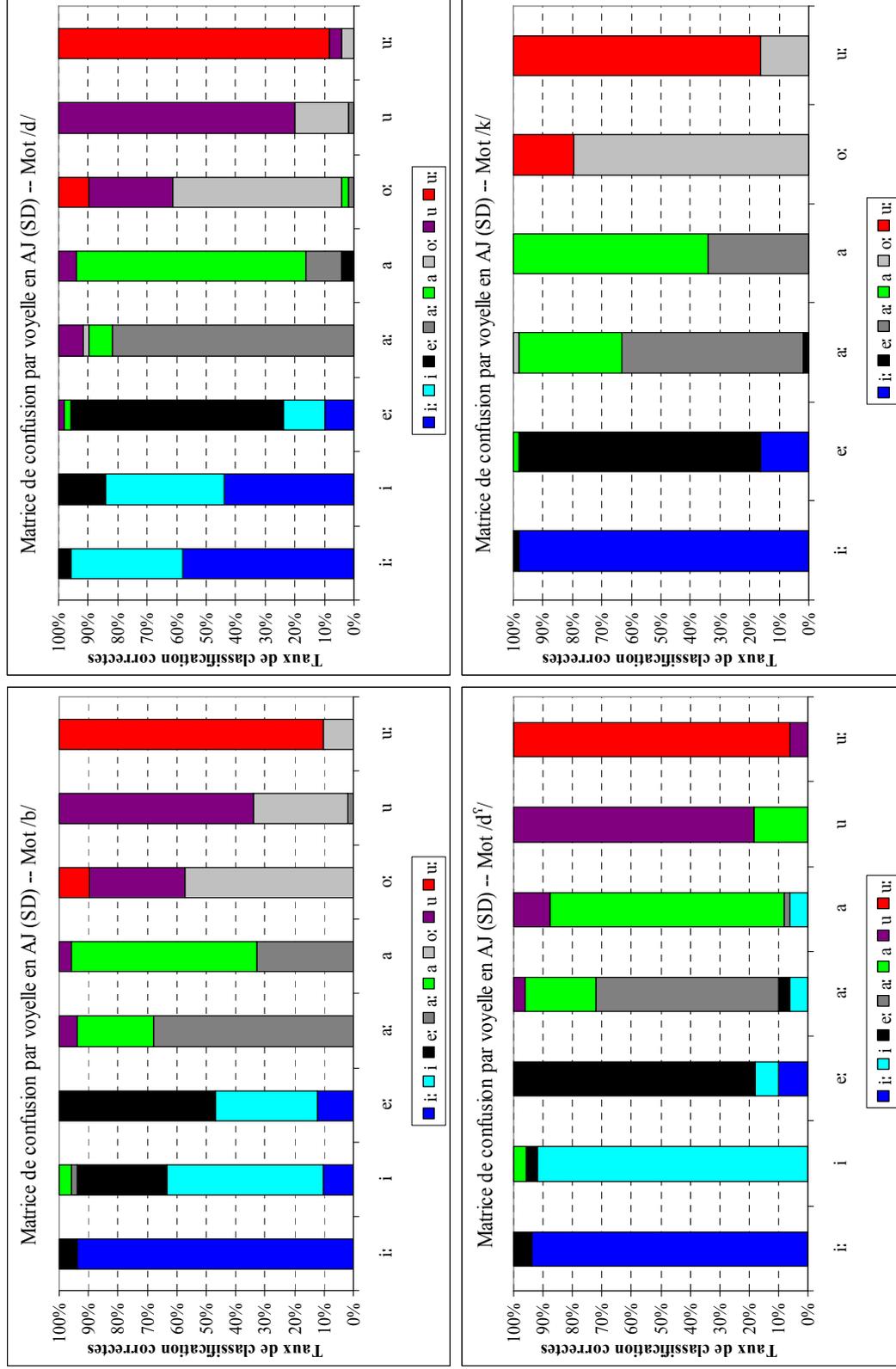


Figure 92 : Matrices de confusion de classification des voyelles de l'AJ en statique sans la durée vocalique (SD) en situation de production Mot et en fonction du lieu d'articulation.

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.

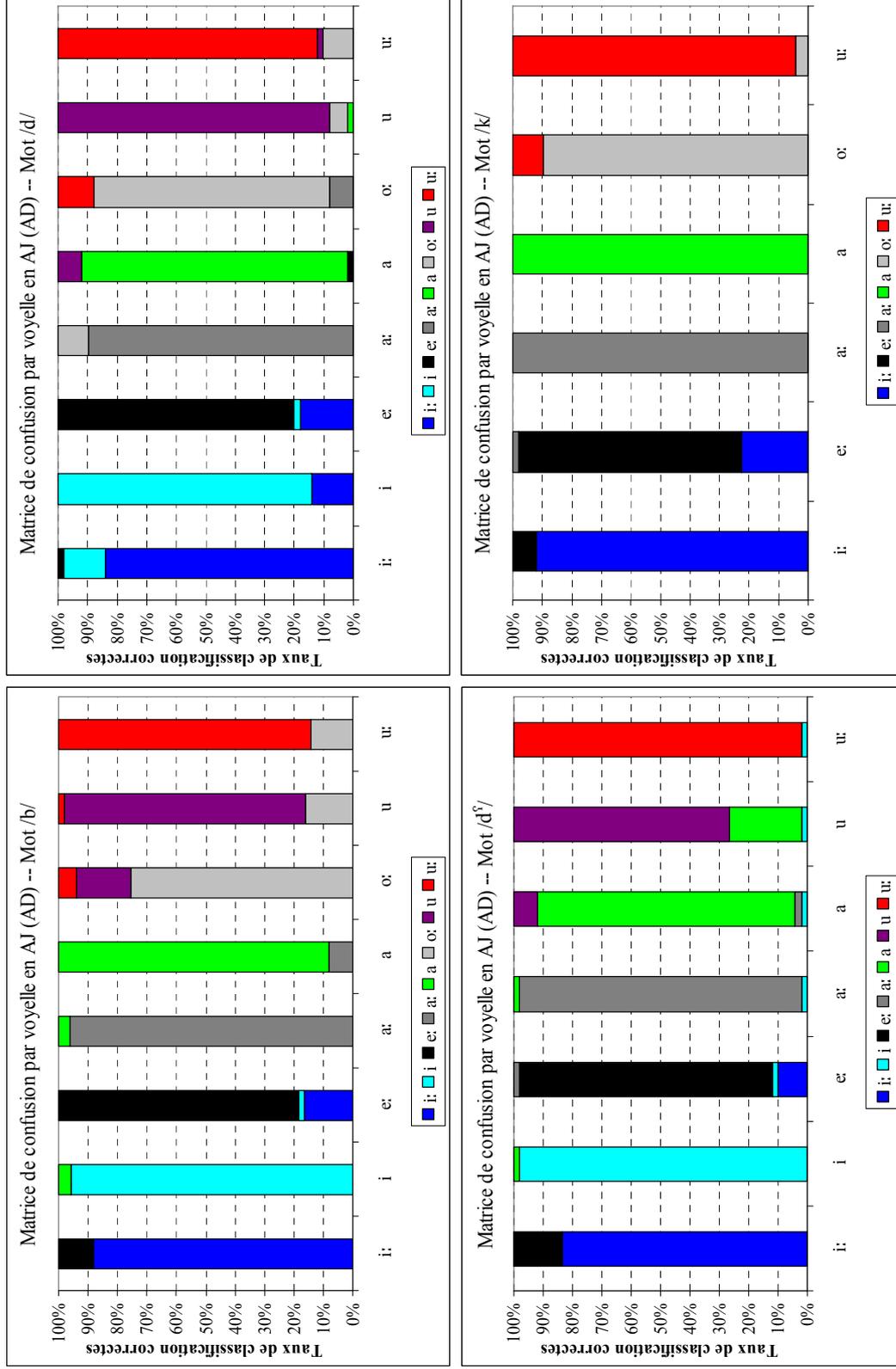


Figure 93 : Matrices de confusion de classification des voyelles de l'AJ en dynamique (régression linéaire) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

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.

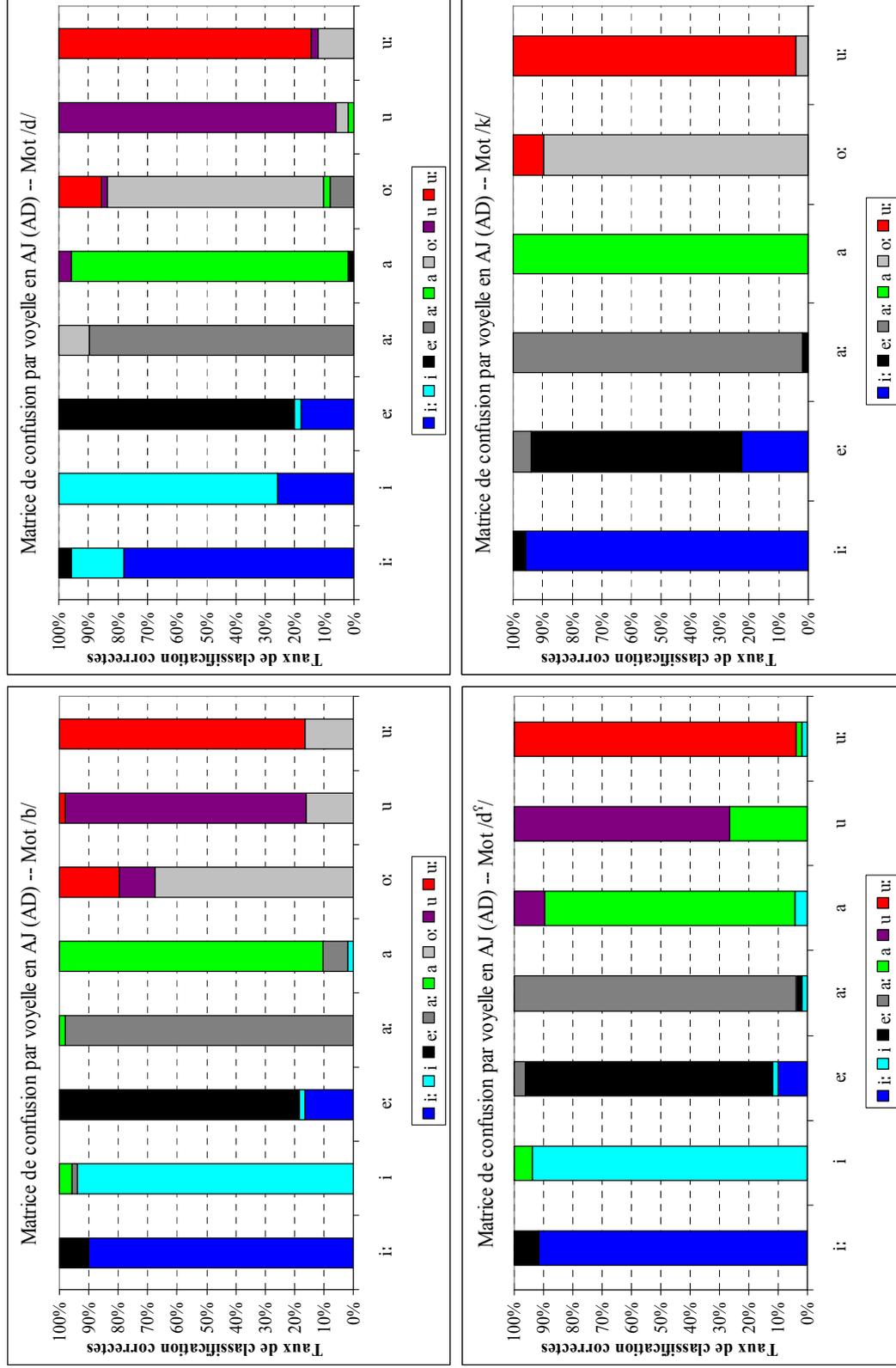


Figure 94 : Matrices de confusion de classification des voyelles de l'AJ en dynamique (régression polynomiale de 2^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

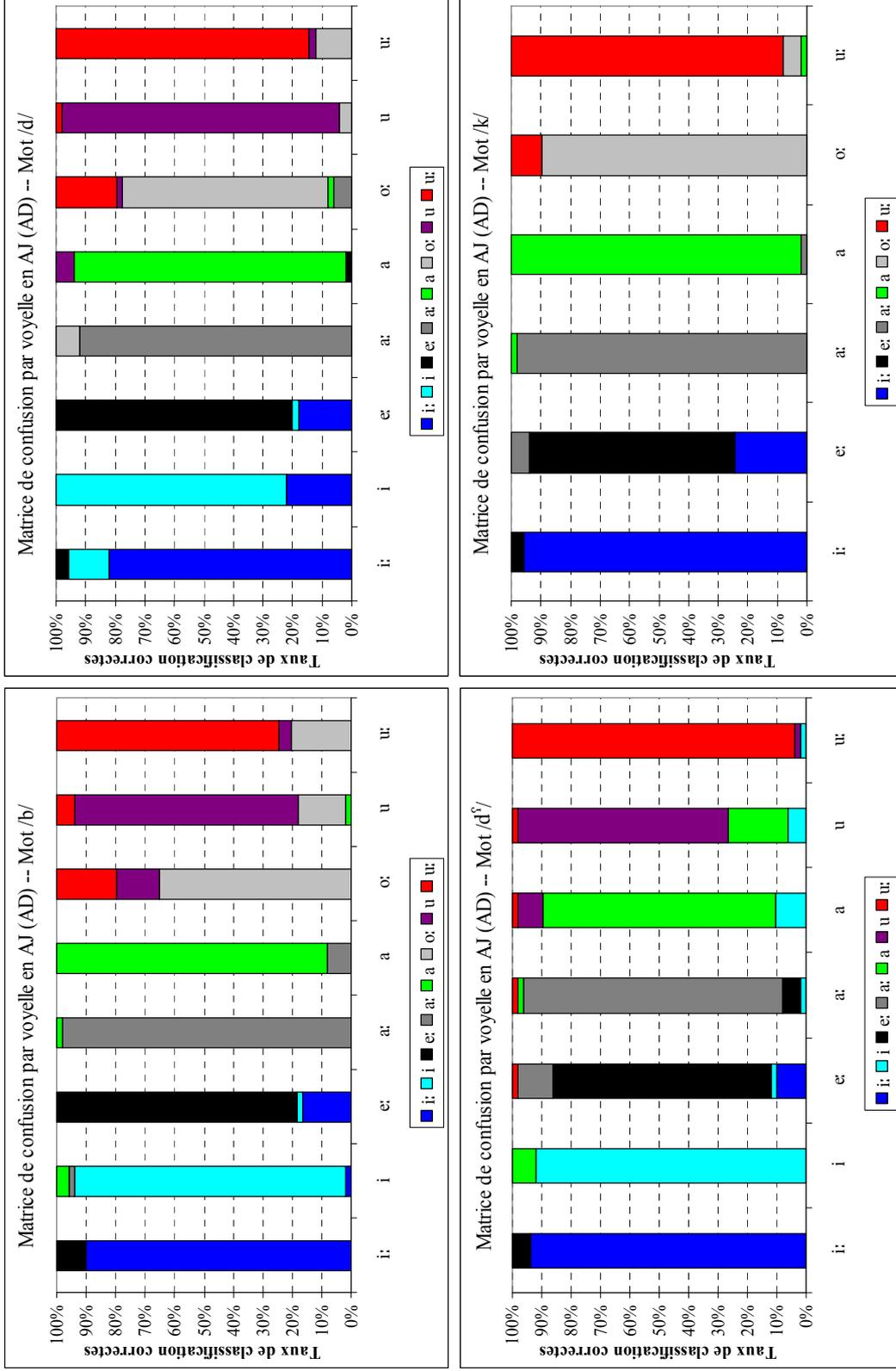


Figure 95 : Matrices de confusion de classification des voyelles de l'AJ en dynamique (régression polynomiale de 3^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

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.

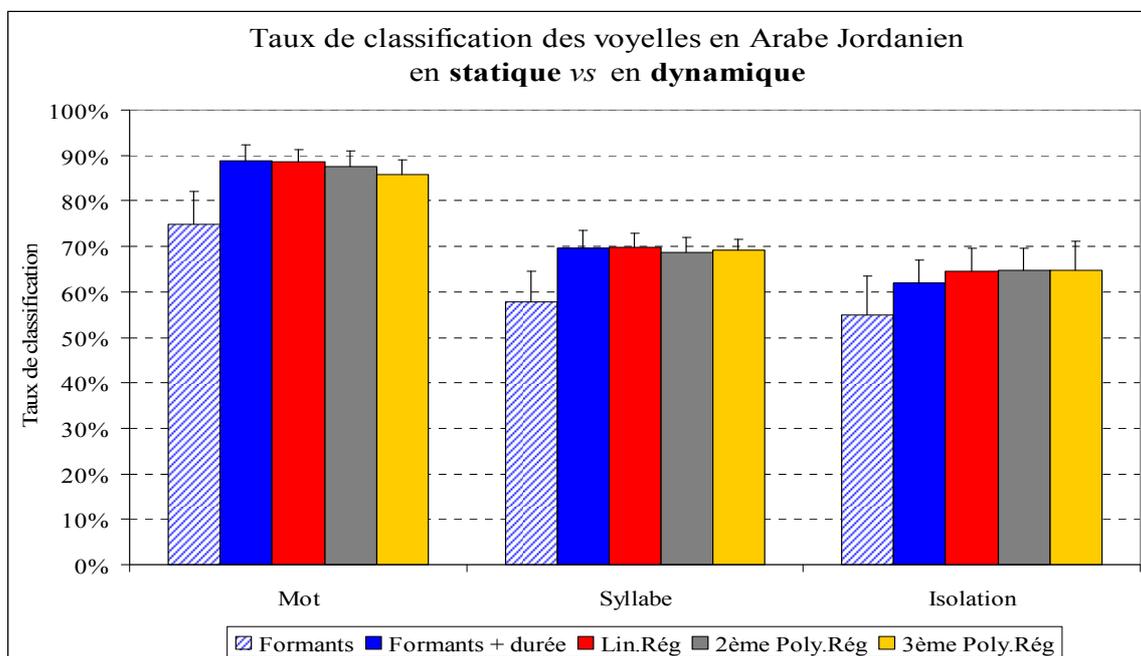


Figure 96 : Taux de classification correcte moyens des voyelles de l'AJ produites dans les contextes /b d d^f k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur).
 Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} 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 ($F(1, 18) = 0,06$; $p = 0,81$)).

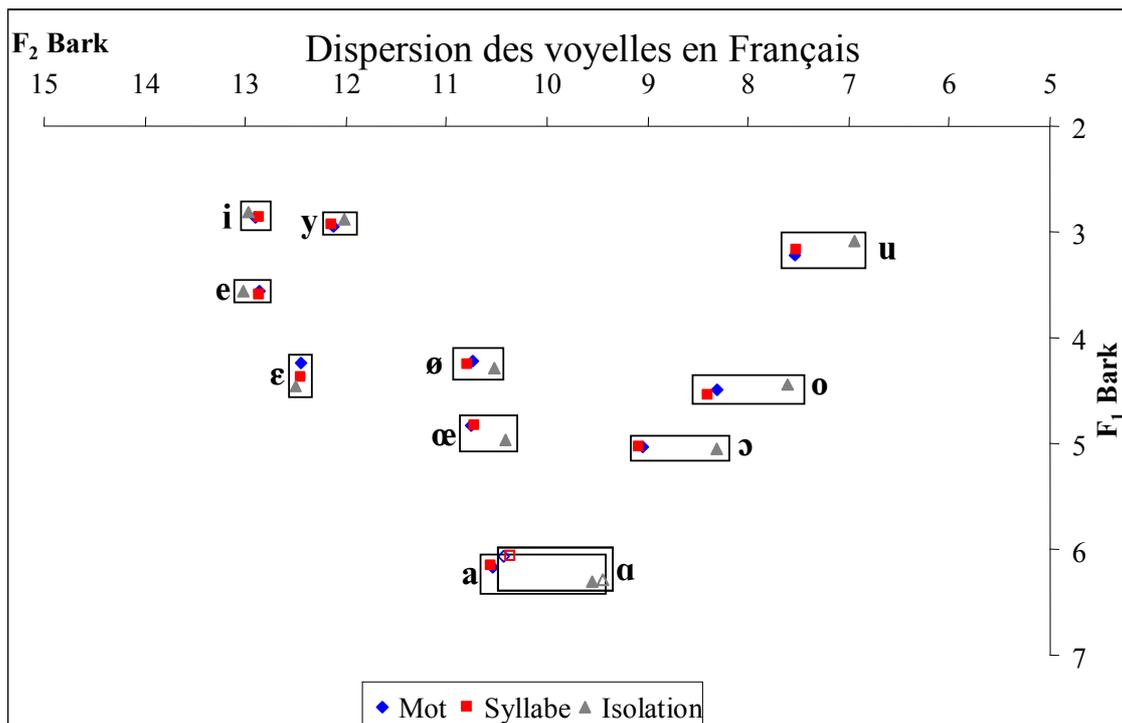


Figure 97 : Espace vocalique en FR en fonction de la situation de production.

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 F₂ ($F(4, 4699) = 7,42 ; p < 0,001$). 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 F₂ uniquement ($F(40, 4699) = 1,86 ; p < 0,001$).

RÉSULTATS EN PRODUCTION DES VOYELLES

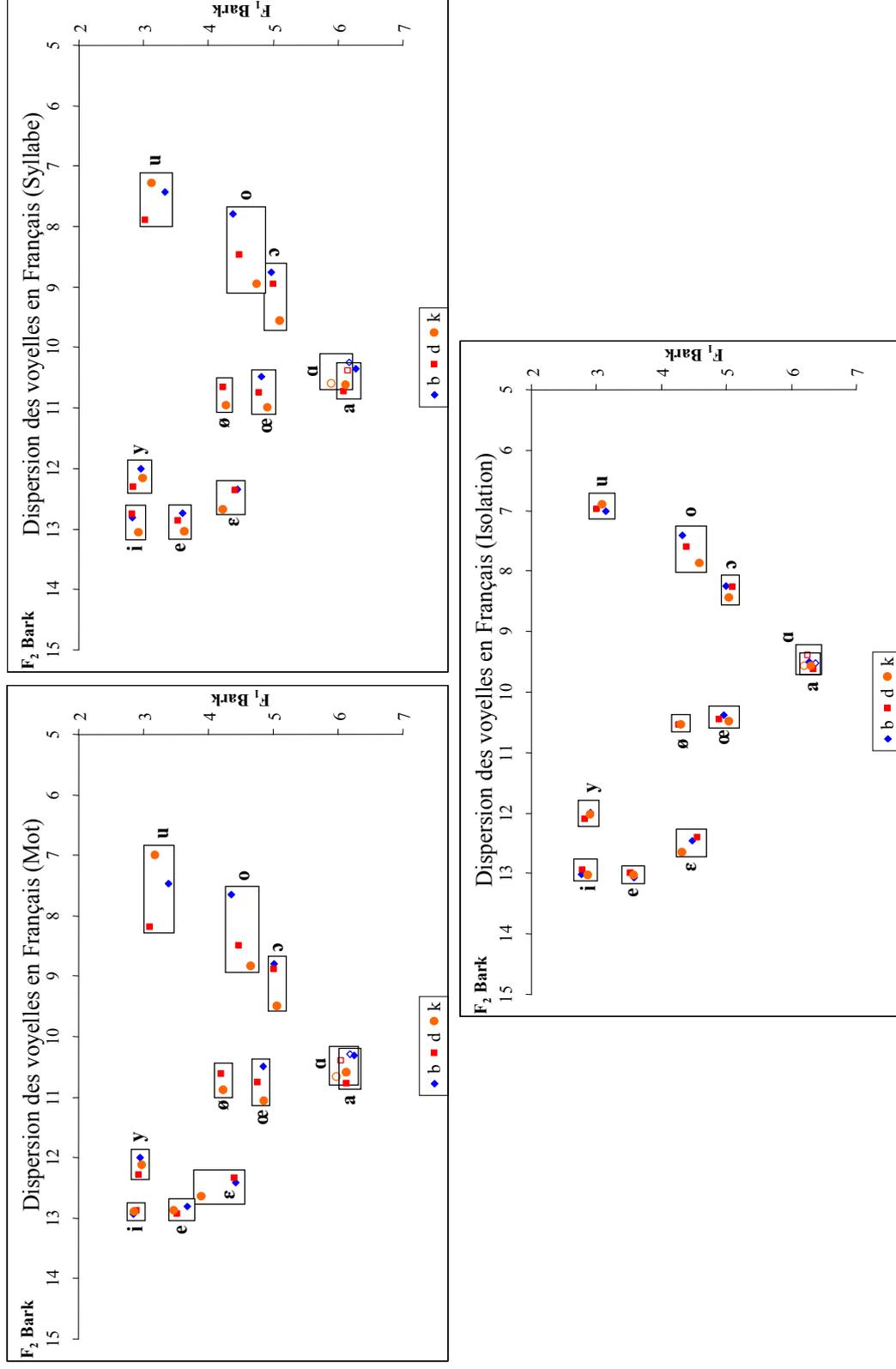


Figure 98 : Espaces vocaliques en FR obtenus en fonction de chaque situation de production et du lieu d'articulation des consonnes /b d k/.

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.

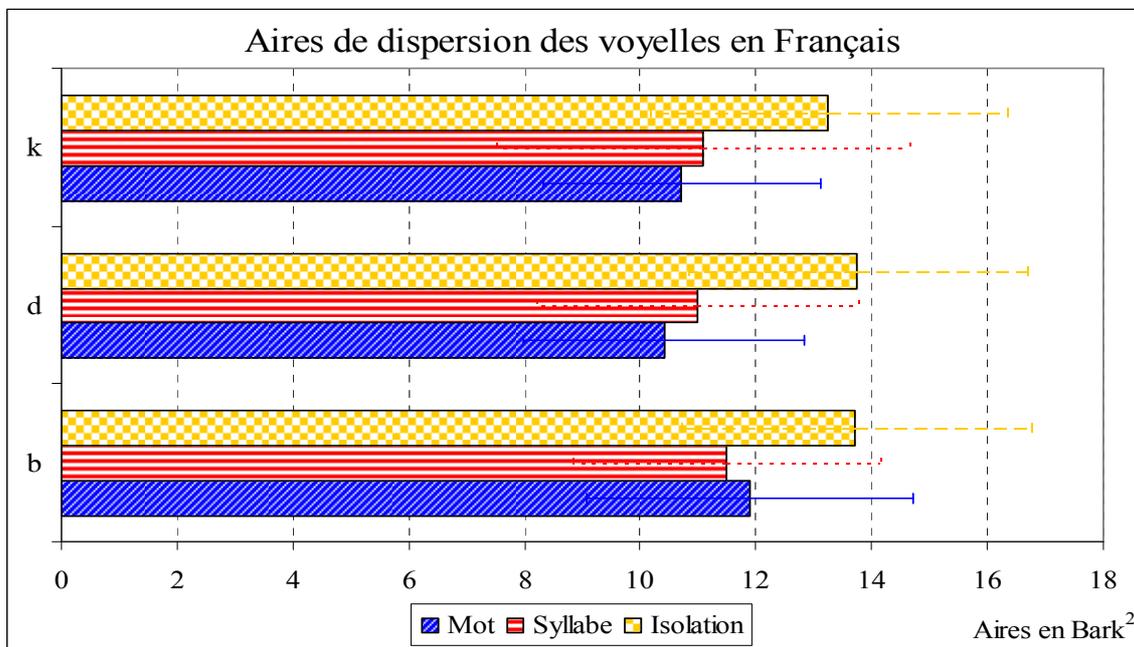


Figure 99 : Aires de dispersion (moyenne et écart-type) des voyelles du FR en fonction de la situation de production et du lieu d'articulation.

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 ».

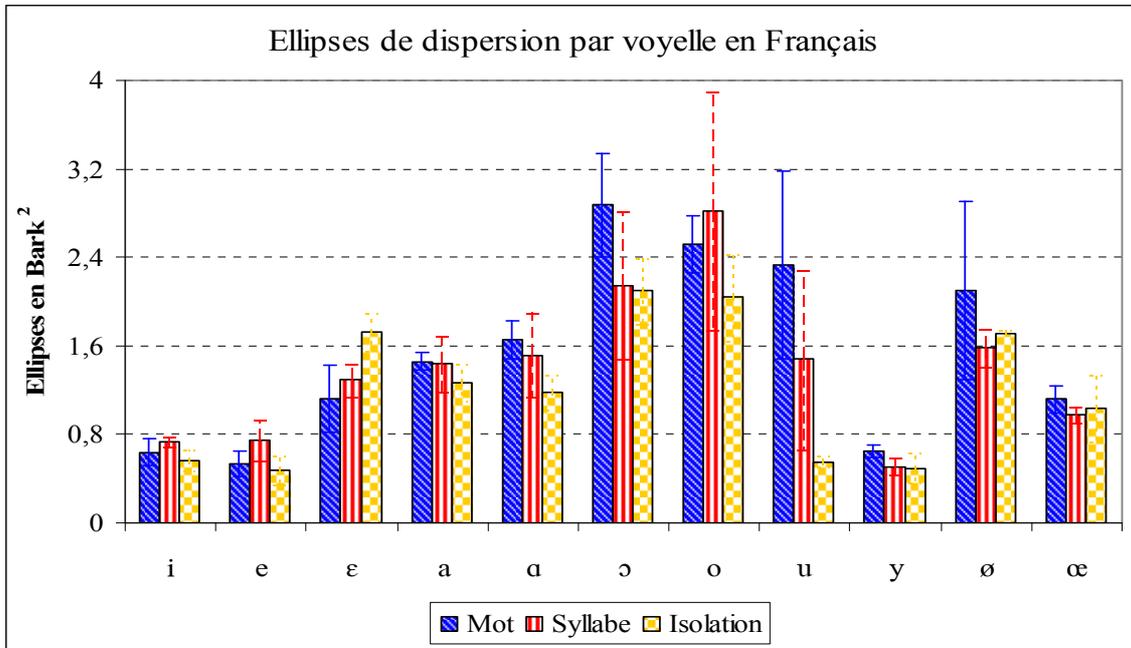


Figure 100 : Taille des ellipses vocaliques en FR en fonction de la situation de production.

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

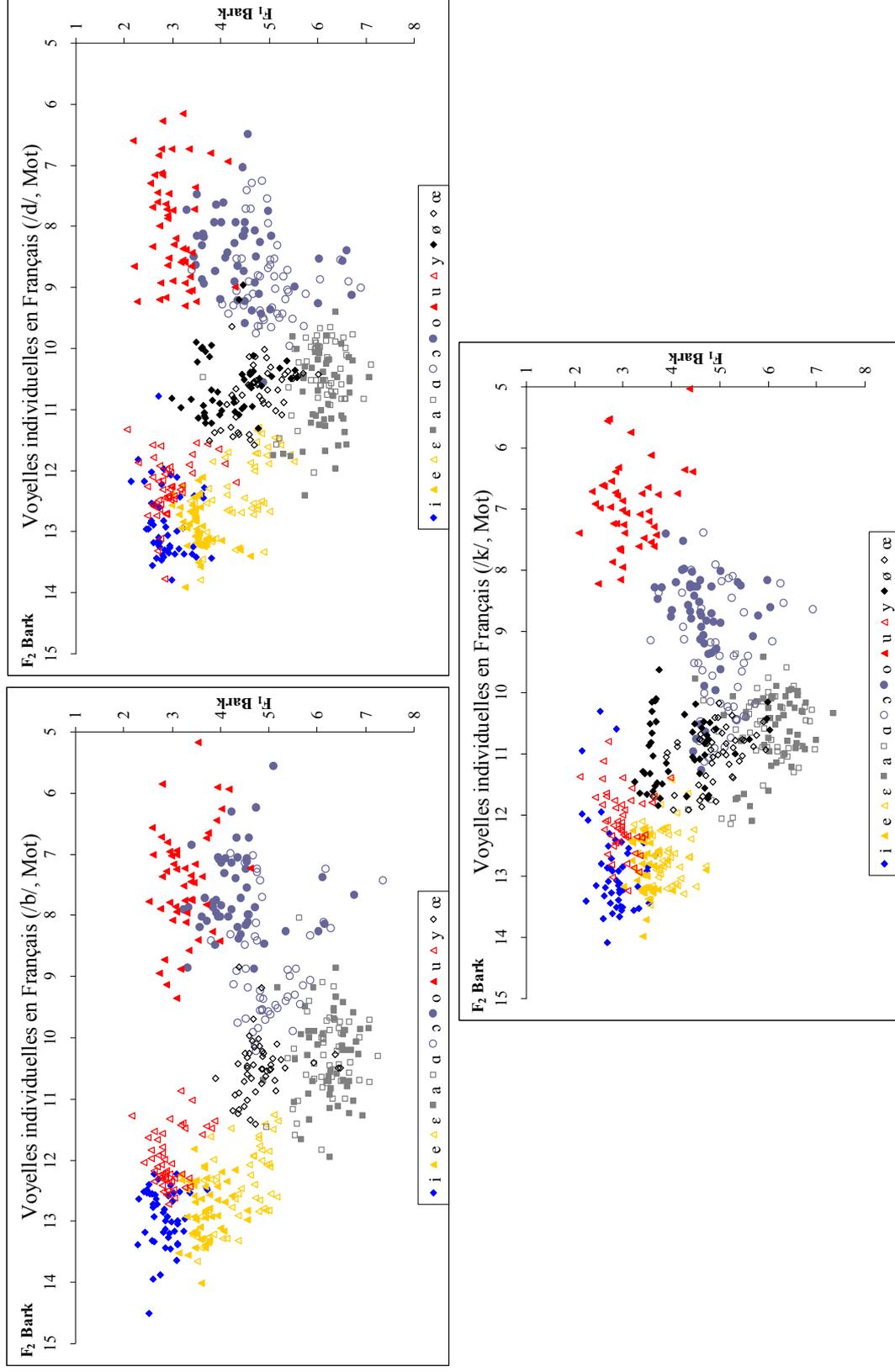


Figure 101 : Dispersion des voyelles en FR à partir des données individuelles, par lieu d'articulation, en Mot.

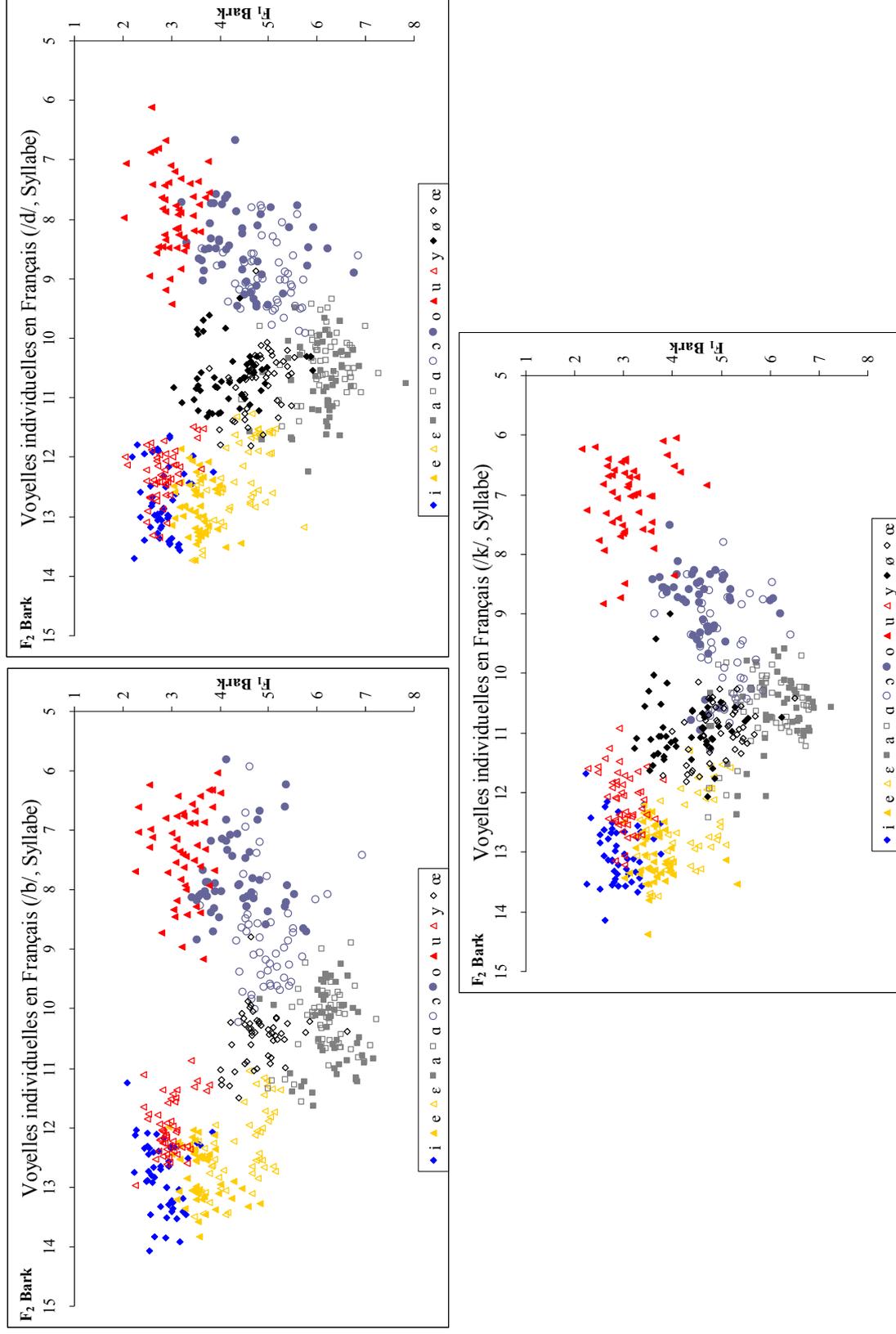


Figure 102 : Dispersion des voyelles en FR à partir des données individuelles, par lieu d'articulation, en Syllabe.

RÉSULTATS EN PRODUCTION DES VOYELLES

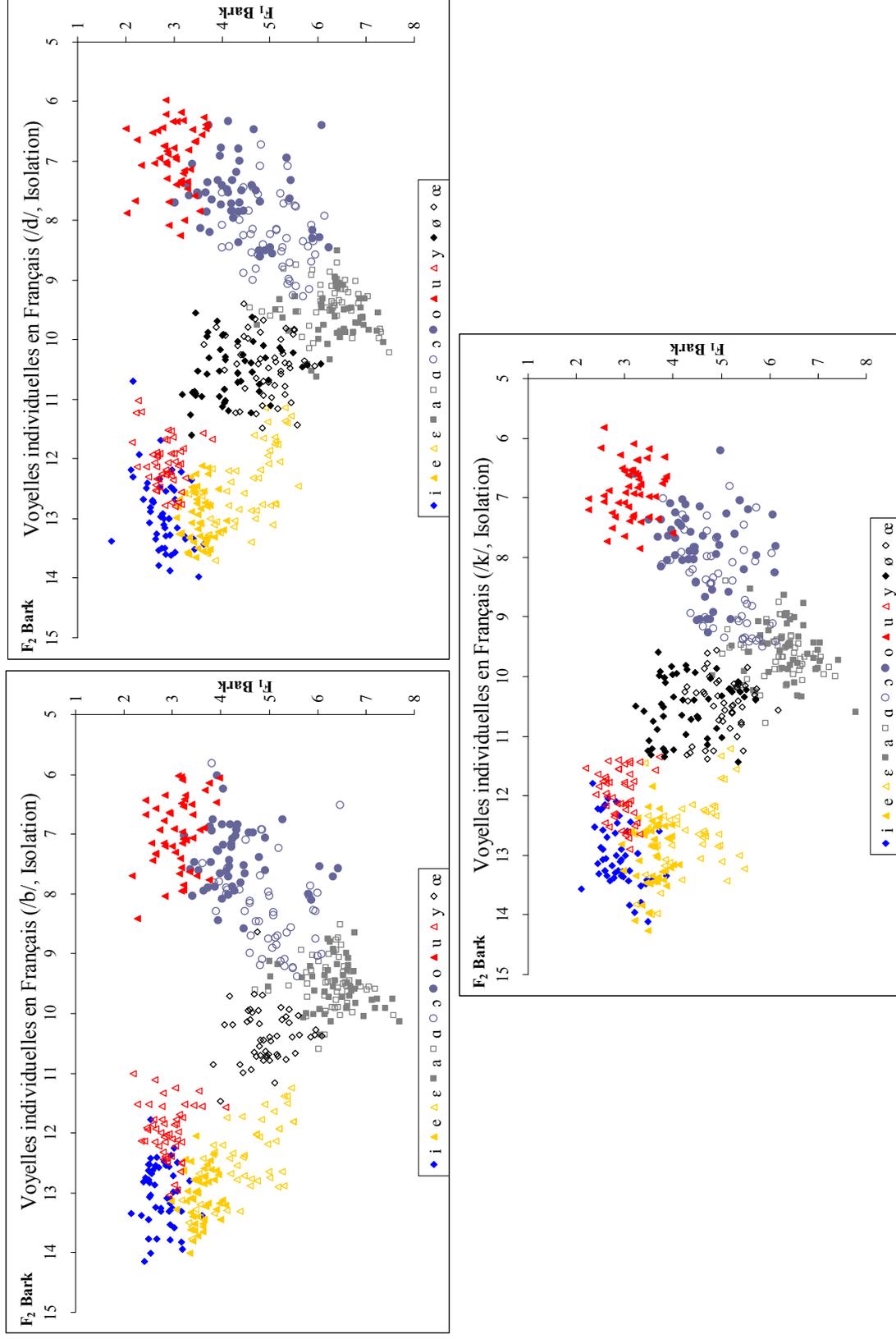


Figure 103 : Dispersion des voyelles en FR à partir des données individuelles, par lieu d'articulation, en Isolation.

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 œ /-/ o ɔ /, 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 / o ɔ /, où le degré d'inclinaison des pentes formantiques est beaucoup plus marqué pour les dernières. En contexte /d/, les quatre voyelles /e \emptyset œ o / présentent le même **onset** vocalique ; les trois voyelles / ϵ œ o / obtiennent le même degré d'inclinaison ; et les couples de voyelles /e ϵ /-/ \emptyset œ /-/ o ɔ / 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 œ /-/ o ɔ /, 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.

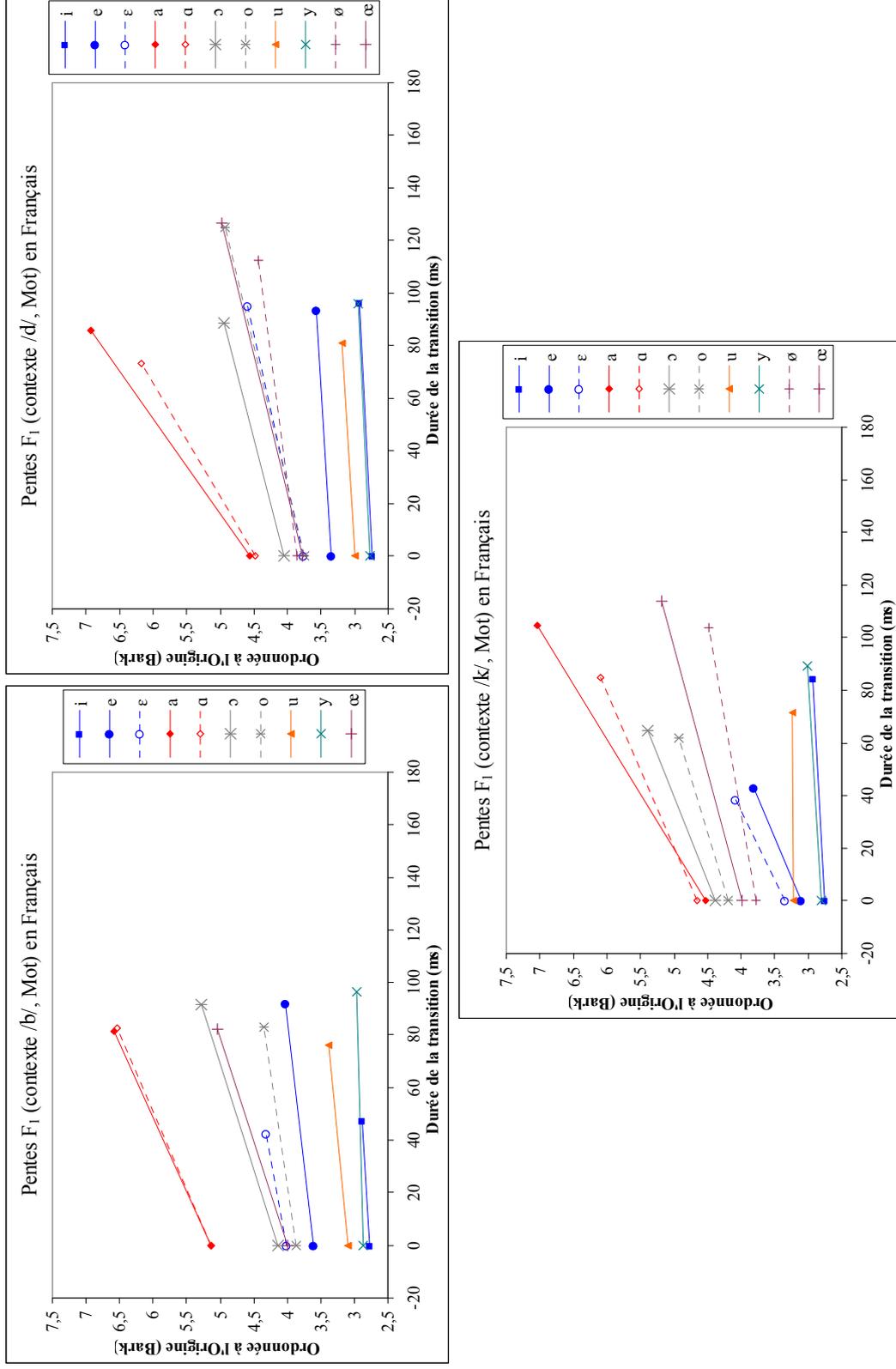


Figure 104 : Pentas formantiques de F₁ des voyelles en FR sans normalisation du temps, en fonction du lieu d'articulation en Mot.

RÉSULTATS EN PRODUCTION DES VOYELLES

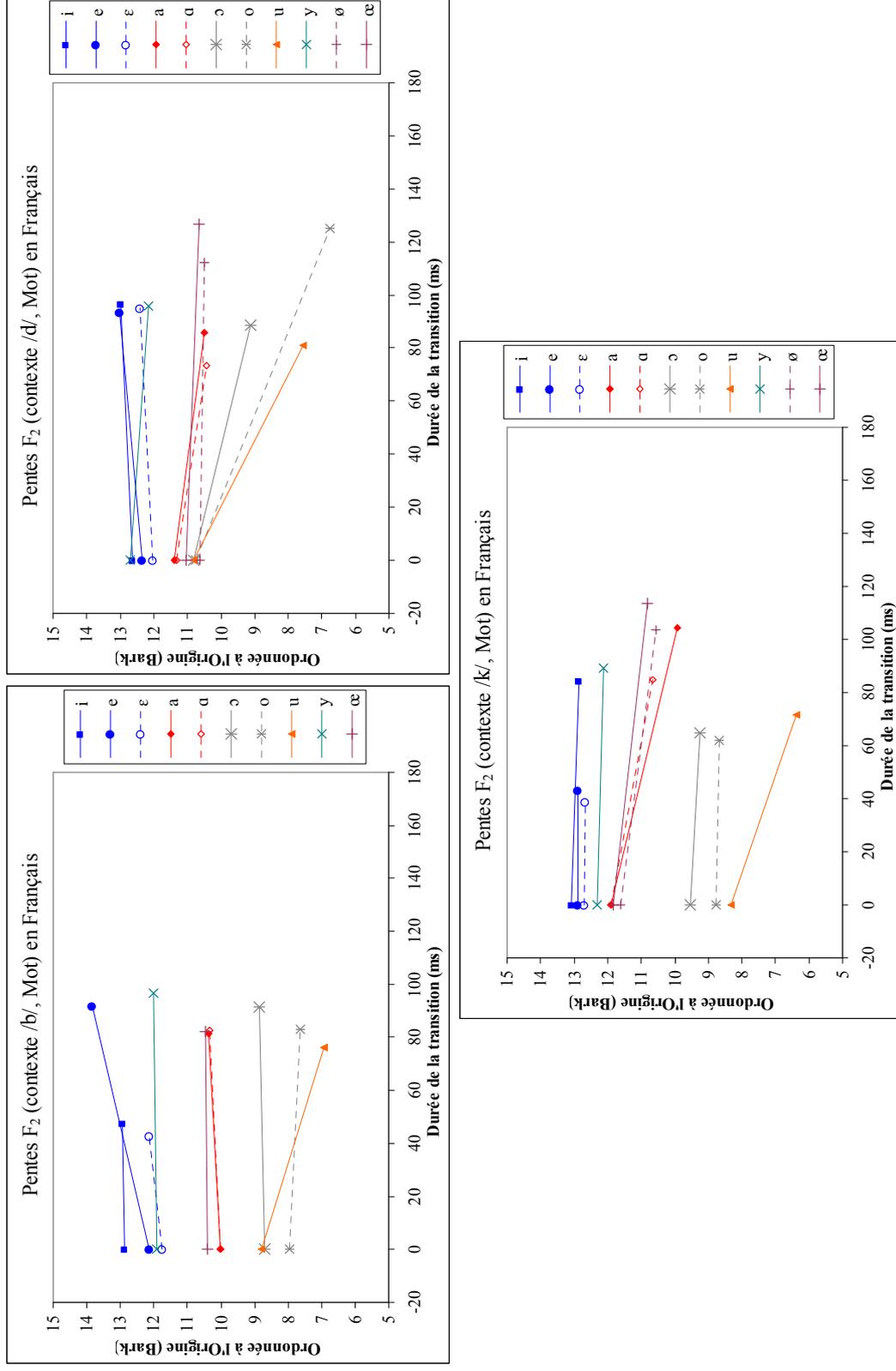


Figure 105 : Pentes formantiques de F₂ des voyelles en FR sans normalisation du temps, en fonction du lieu d'articulation en Mot.

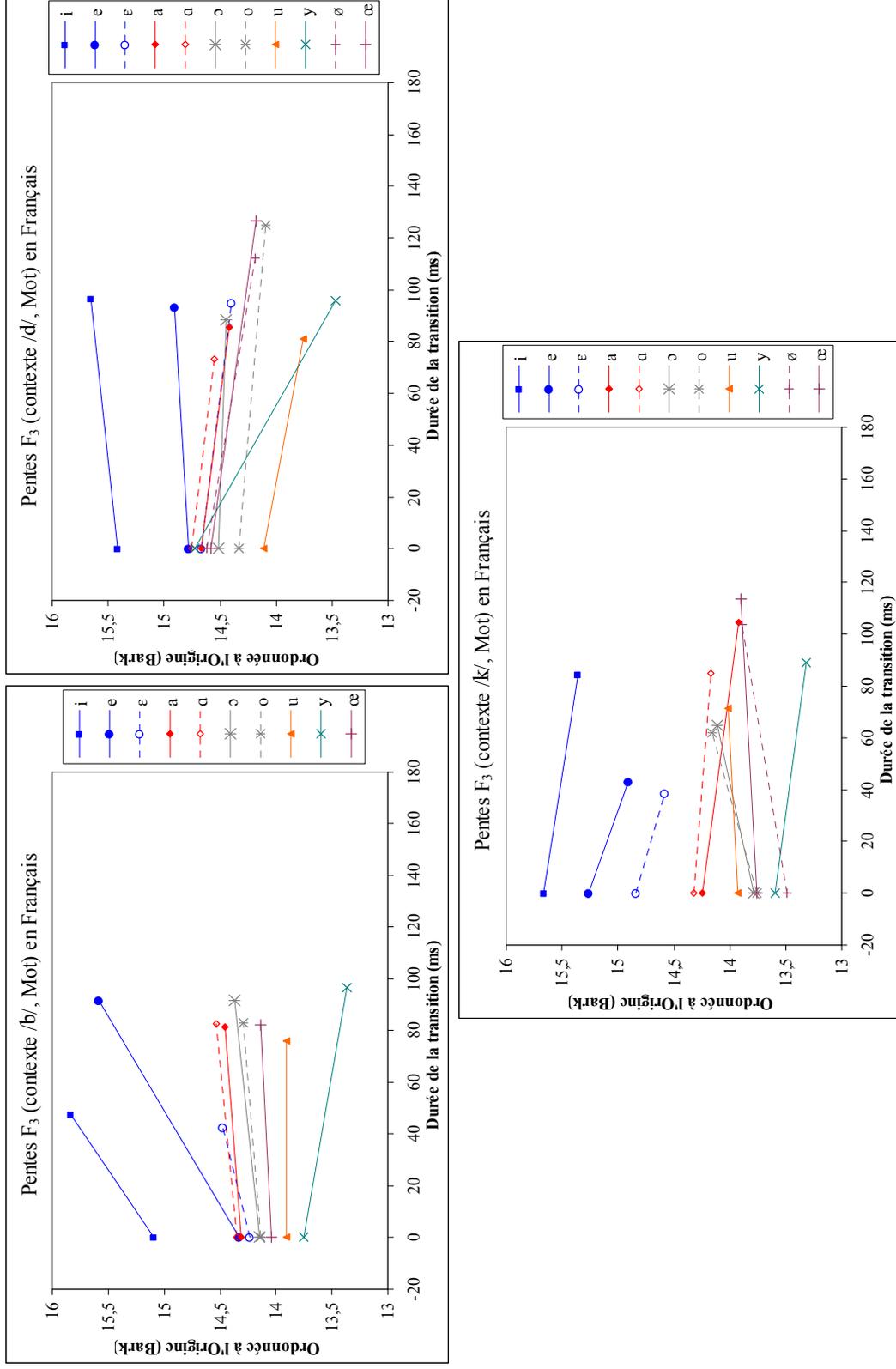


Figure 106 : Pentas formantiques de F₃ des voyelles en FR sans normalisation du temps, en fonction du lieu d'articulation en Mot.

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**.

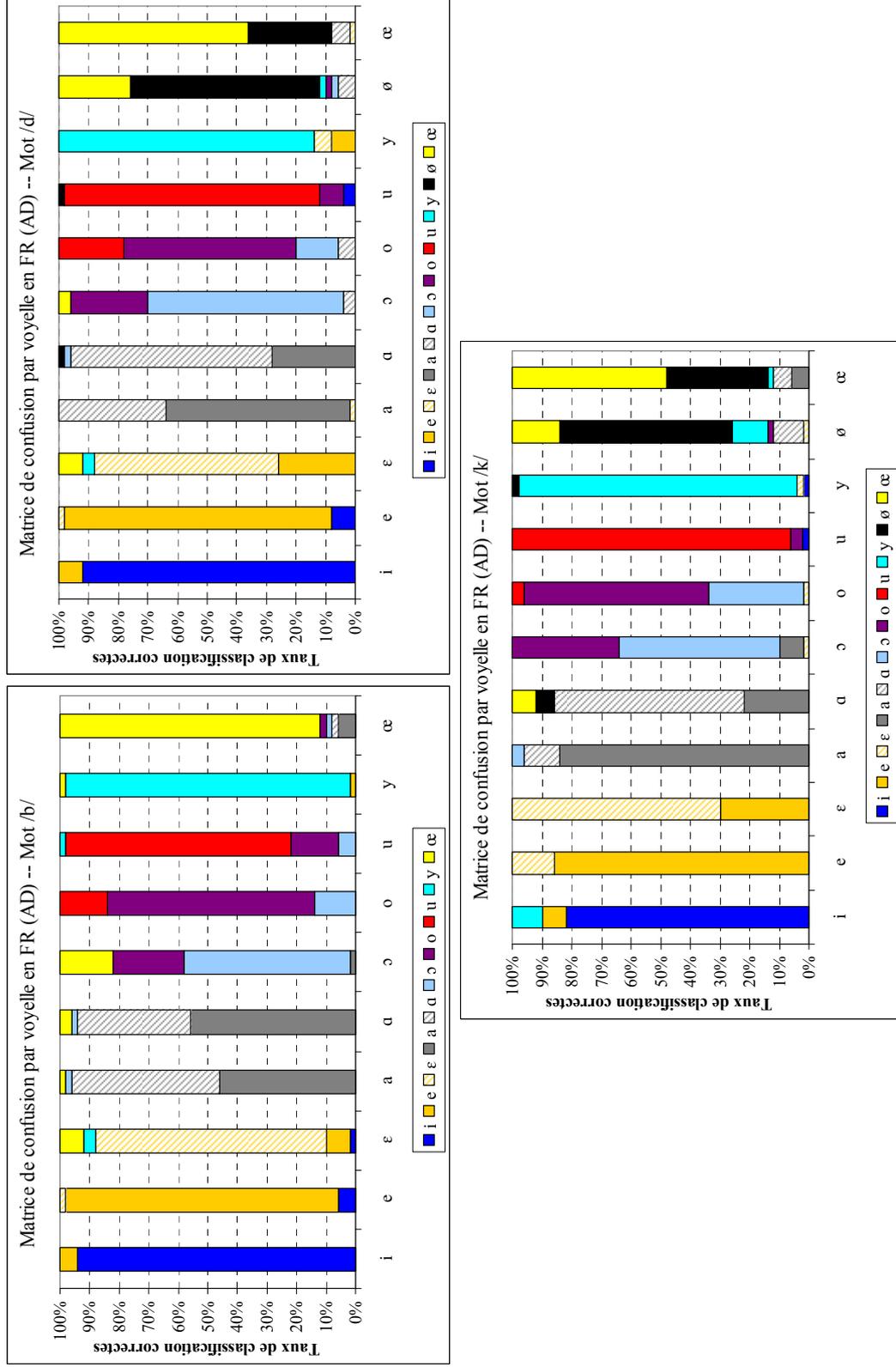


Figure 107 : Matrices de confusion de classification des voyelles du FR en **statique** avec la durée vocalique (AD) en situation de production Mot et en fonction du lieu d'articulation.

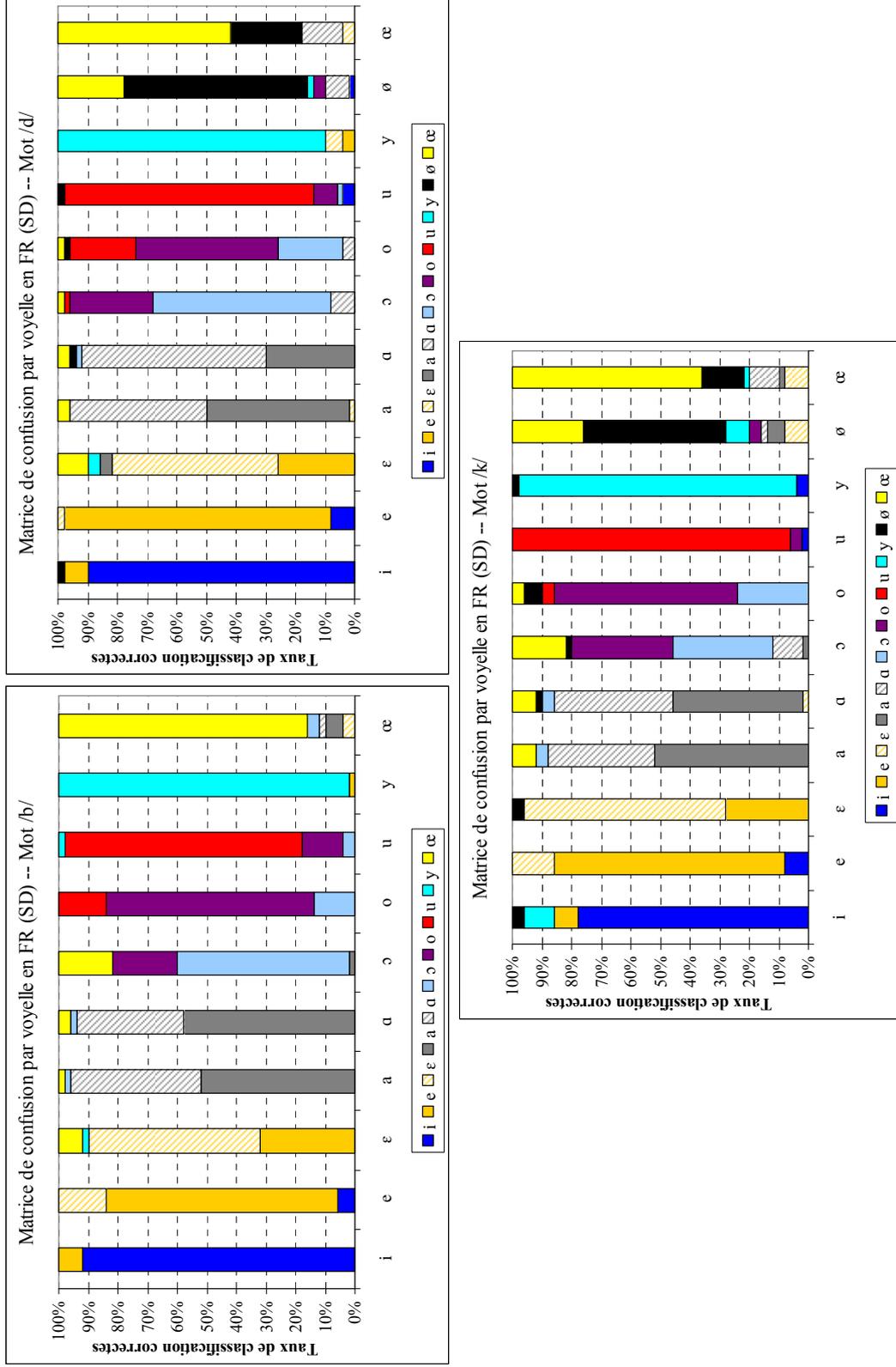


Figure 108 : Matrices de confusion de classification des voyelles du FR en statique sans la durée vocalique (SD) en situation de production Mot et en fonction du lieu d'articulation.

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.

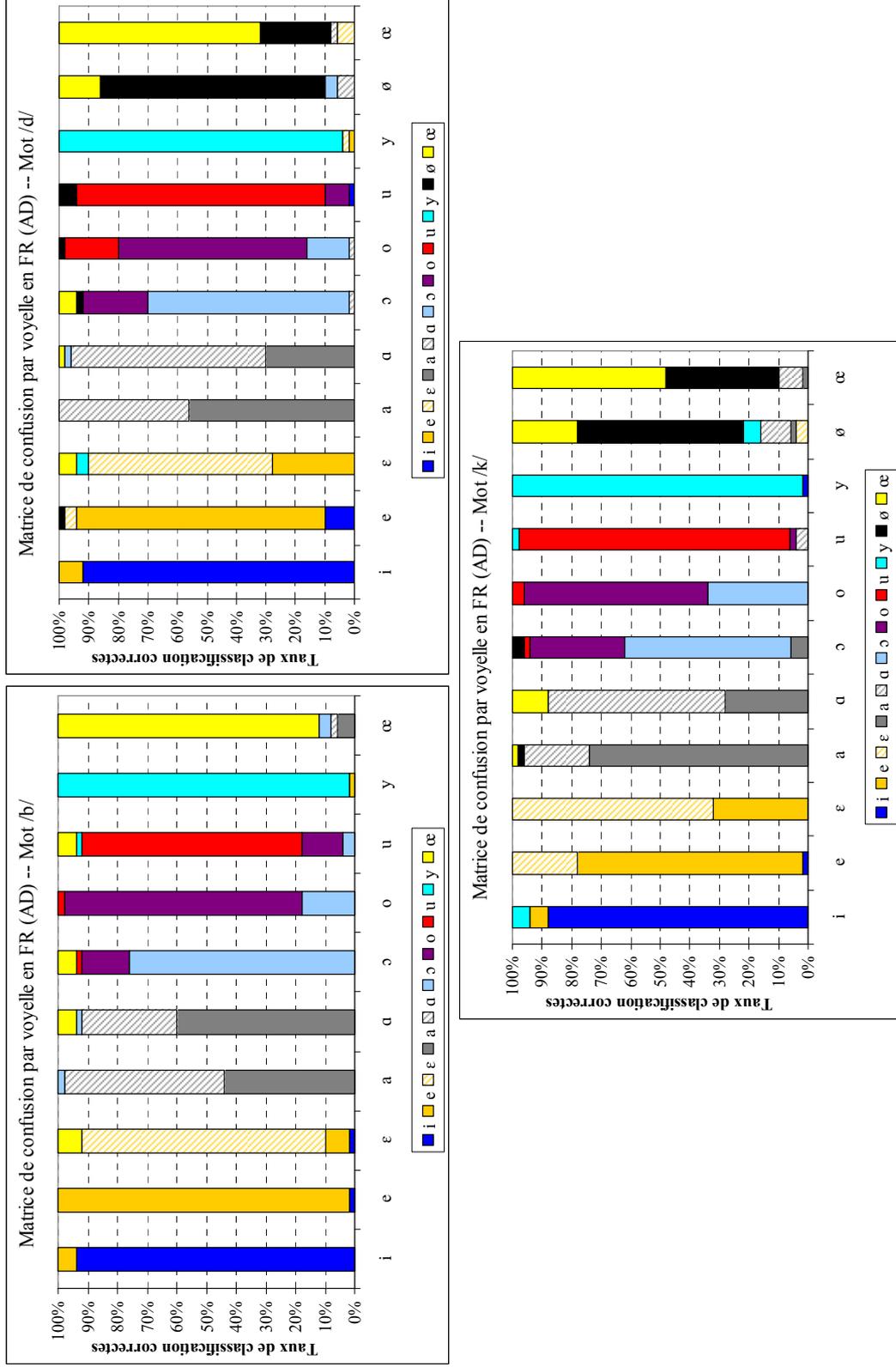


Figure 109 : Matrices de confusion de classification des voyelles du FR en dynamique (régression linéaire) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

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%.

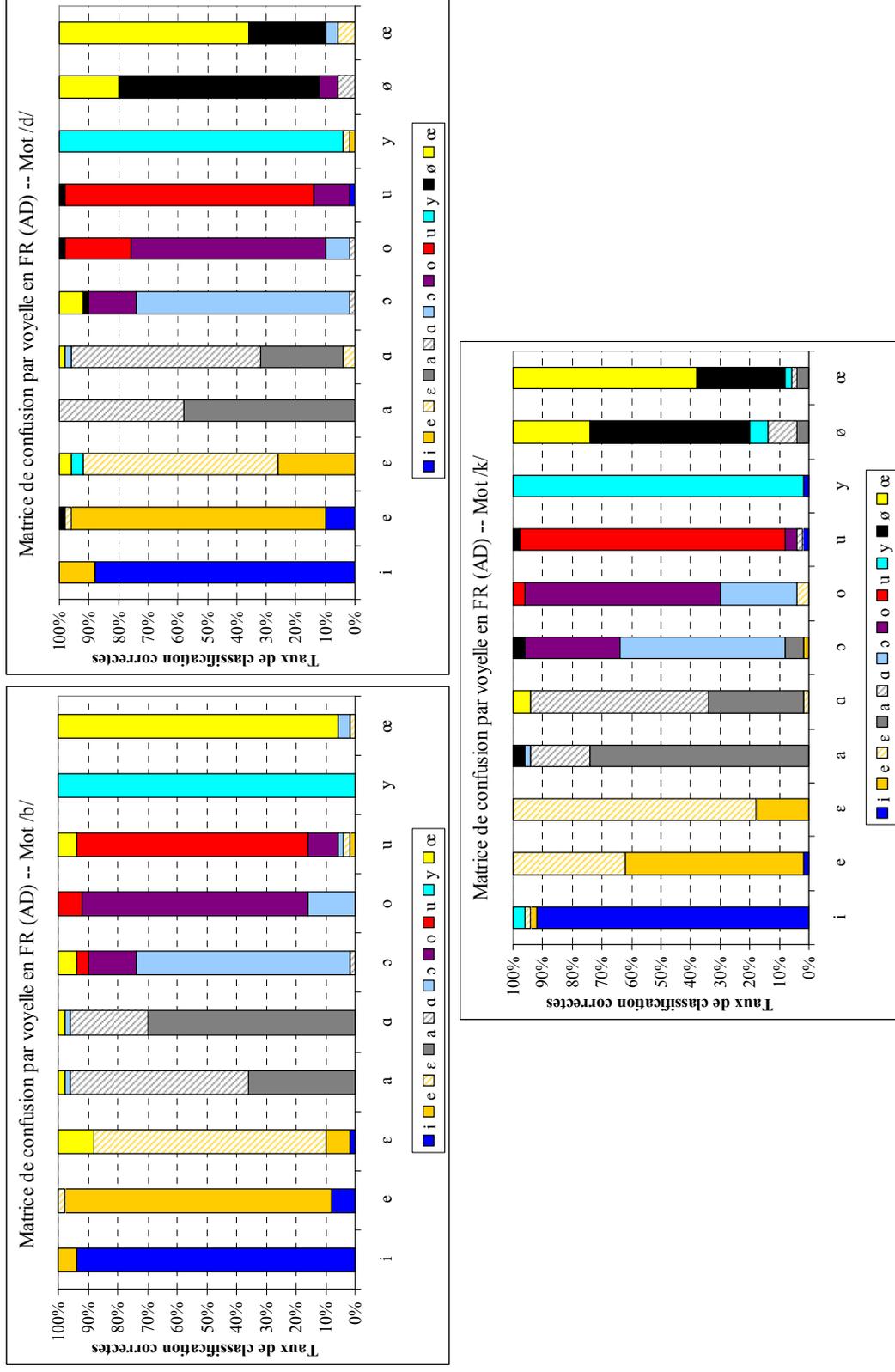


Figure 110 : Matrices de confusion de classification des voyelles du FR en dynamique (régression polynomiale de 2^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

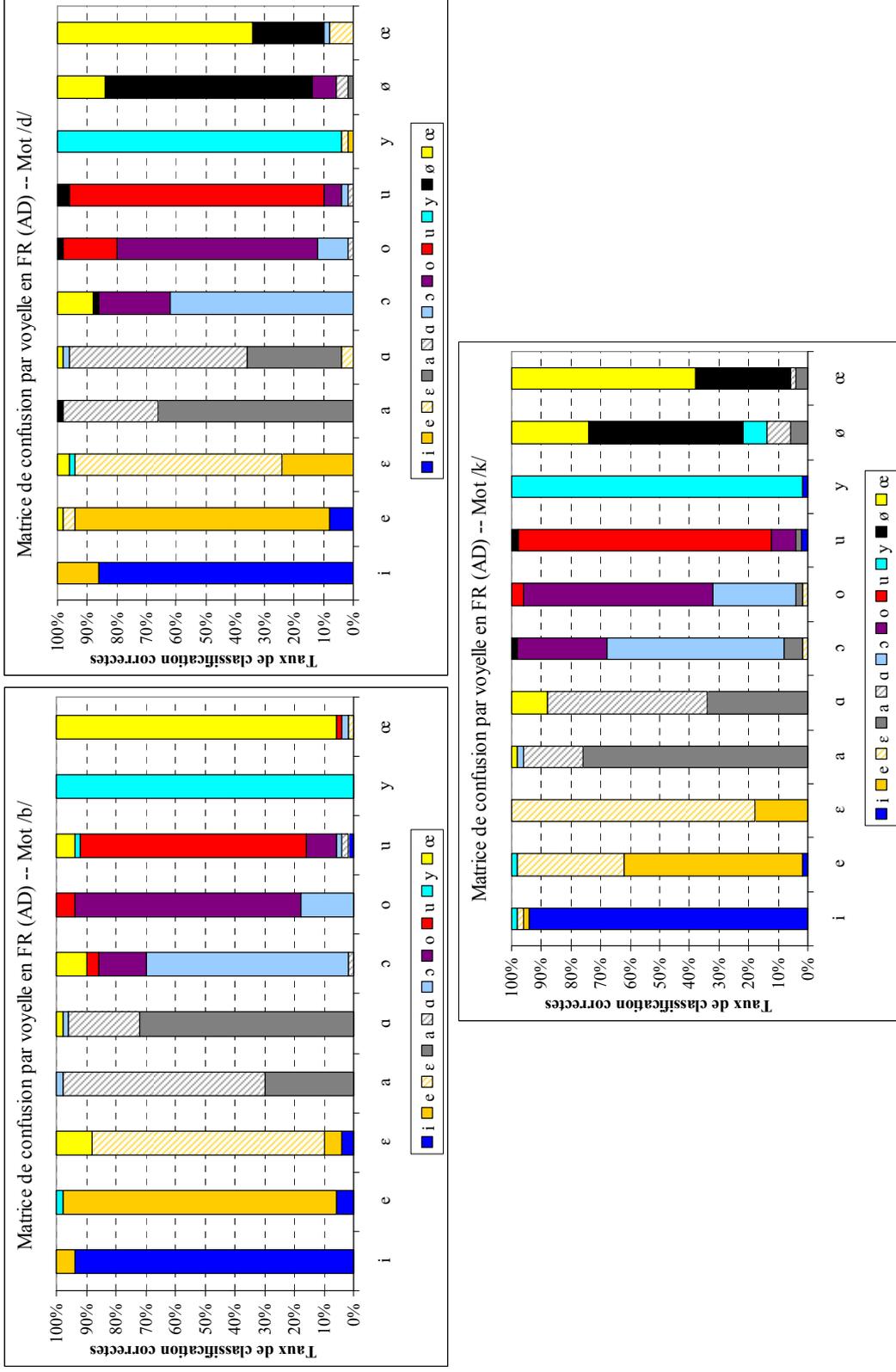


Figure 111 : Matrices de confusion de classification des voyelles du FR en **dynamique** (régression polynomiale de 3^{ème} ordre) avec la durée de la transition formantique (AD) en situation de production Mot et en fonction du lieu d'articulation.

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.

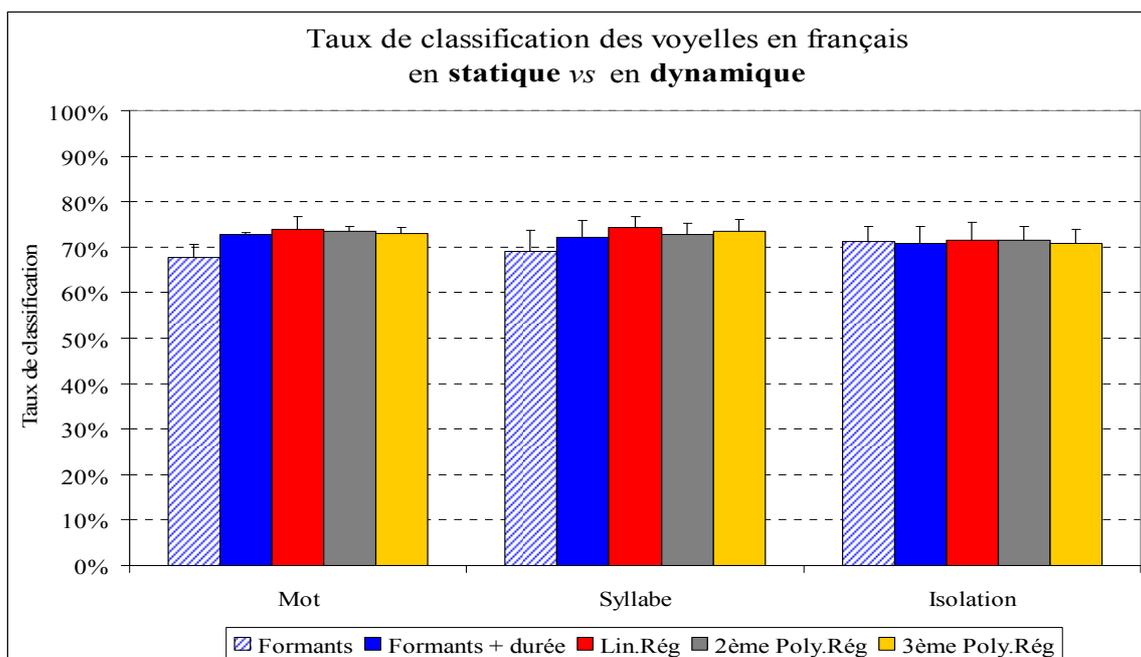


Figure 112 : Taux de classification correcte moyens des voyelles du FR produites dans les contextes /b d k/ en fonction des trois situations de production, en **statique** (fond hachuré) vs en **dynamique** (fond en couleur).
 Lin.Rég = régression linéaire, Poly.Rég = régression polynomiale (2^{ème} 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.