

## Chapitre 6

# Un modèle préliminaire des changements linguistiques

*Une somme d'expériences ne peut jamais me prouver que j'ai raison ;  
une seule expérience peut n'importe quand me prouver que je me suis trompé.*  
Albert Einstein.

Nous avons tenté dans le deuxième chapitre de cet ouvrage d'offrir une vision générale des changements linguistiques. Pour cela, nous avons d'une part abordé les concepts de la sociolinguistique pour rendre compte du problème de l'implémentation du changement dans une communauté linguistique et du poids des liens sociaux. D'autre part, nous avons insisté sur les aspects internes et systémiques des systèmes linguistiques, dans la lignée de l'approche structuraliste Saussurienne. Ces phénomènes jouent également un rôle dans l'actuation et dans l'implémentation d'un changement. Toutefois, aujourd'hui, si les travaux portant sur l'un ou l'autre de ces aspects sont nombreux, l'étude de l'intersection des contraintes sociales et des contraintes systémiques des langues demeure un problème difficile, de part la difficulté de pouvoir étudier "naturellement" (en situation) l'interaction des structures linguistiques dans un contexte social.

Nous souhaitons aborder ici ce problème de façon très modeste. Dans la lignée des travaux de Nettle [Nettle, 1999b] [Nettle, 1999c] [Nettle, 1999a], nous pensons que les modèles informatiques constituent un outil utile pour dégager les tenants et les aboutissants de ces phénomènes.

A cette fin, nous présentons tout d'abord un modèle des changements linguistiques, que nous détaillons selon ces composantes internes et sociales (externes).

Ceci fait, dans une seconde partie, nous présentons différentes expériences informatiques afin de dégager quelques implications de notre modèle, et voir si celles-ci ont une pertinence vis à vis des changements réels des langues du monde. Nous examinons ainsi successivement la dynamique d'évolution d'un unique système soumis à un certain jeu de contraintes naturelles et distribuées, puis l'impact des interactions sociales entre les individus. Dans le cadre de notre questionnement sur l'origine et le développement des langues, notre dernier objectif est alors de mettre à profit les résultats précédents pour contraster même très partiellement les évolutions des langues contemporaines et celles des langues de la préhistoire. L'influence des structures démographiques décrites au chapitre 4 sera au cœur de nos interrogations.

## 6.1 Description du modèle

### 6.1.1 Introduction à la notion d'optimisation et aux techniques mathématiques de minimisation de fonction

Avant de rentrer précisément dans le modèle qui nous intéressera dans ce chapitre, il est nécessaire d'introduire rapidement quelques concepts mathématiques liés à l'optimisation de fonction. Ces notions seront en effet utiles pour comprendre les différentes dynamiques d'évolution des systèmes linguistiques que nous introduirons par la suite.

#### Notion d'optimisation de fonction et paysages énergétiques

Le domaine mathématique des techniques d'optimisation a pour objet la recherche des optima de fonctions de une ou plusieurs variables de la façon la plus efficace qui soit. L'optimisation de fonction joue un rôle important dans le domaine de l'analyse numérique, qui concerne des problèmes mathématiques qu'il n'est pas possible de traiter de façon analytique, et qui requièrent donc des approches numériques pour être résolus<sup>39</sup>.

Nous avons rapidement présenté au chapitre 2 la notion de paysage énergétique (*fitness landscape*) et l'existence dans ces paysages de puits qui correspondent à des minima énergétiques. Le plus souvent, le but d'une étude est la découverte des puits de plus grande profondeur, qui correspondent à des états particuliers et "intéressants" du système, par exemple en termes de performances pour un ingénieur. L'optimisation correspond le plus souvent à une minimisation de fonction (bien que minimisation ou maximisation soient des opérations totalement symétriques), et la fonction à minimiser est le plus souvent celle qui permet le calcul de l'énergie du système à partir des paramètres qui régissent ce dernier.

Nous pouvons citer comme exemple très actuel la découverte des minima énergétiques dans l'espace des conformations de molécules, qui correspondent aux formes que pourra prendre une molécule en réponse à ses contraintes internes.

Découvrir un minimum d'énergie (un puits) est une tâche relativement difficile puisque la topographie du paysage est généralement inconnue et qu'un nombre de paramètres un tant soit peu élevé conduit à un immense domaine dont l'exploration exhaustive à la recherche de minima énergétiques est computationnellement trop difficile.

Différentes techniques sont envisageables pour rechercher les minimaux énergétiques (voir par exemple [Ciarlet, 1982]). Une classification de ces méthodes peut se faire selon leur ordre, c'est à dire selon leur utilisation des dérivées de la fonction de calcul d'énergie. Une méthode d'ordre 0 consiste simplement à calculer la valeur de l'énergie en des points formant un maillage de l'espace des paramètres, et à choisir le point d'énergie minimale. Bien sûr, plus la densité du maillage sera fine, plus la valeur trouvée sera proche de l'optimum réel. Encore une fois, le problème est qu'un grand nombre de paramètres conduira à un nombre gigantesque de points pour un maillage de forte densité, forte densité nécessaire pour surmonter les possibles variations abruptes de la fonction selon certains paramètres.

D'autres méthodes d'ordre 0 peuvent emprunter à ce qu'il est convenu d'appeler les méthodes de Monte-Carlo (sous ce terme se cachent un grand nombre de méthodes dans de multiples dis-

---

<sup>39</sup>C'est le cas actuellement de très nombreuses équations aux dérivées partielles, que ce soit en physique, en chimie...

ciplines), qui consistent à effectuer des échantillonnages stochastiques de l'espace énergétique pour tenter de découvrir les minima à l'aide de méthodes statistiques.

Les méthodes d'ordre 1, comme la méthode de plus grande pente (*steepest descent*) ou celle des gradients conjugués (*conjugated gradients*) mettent à profit la dérivée première de la fonction énergétique. Les méthodes d'ordre 2, comme la minimisation de Newton-Raphson, prennent en compte la courbure de la fonction d'énergie pour converger encore plus vite vers un minimum. Dans tous les cas, le principe est de partir d'une position initiale dans l'espace des paramètres, et d'envisager une évolution du système (par modification des paramètres) qui suive les structures du relief vers les points de plus faible altitude (plus la méthode est d'ordre supérieur, plus le relief est suivi "intelligemment" et rapidement vers le minimum). Arrivé à un minimum (dérivée de la fonction d'énergie nulle, courbure positive), le système se stabilise.

Le principal problème dans la recherche des minima absolus (c'est à dire les puits les plus profonds) est la possibilité que les méthodes précédentes (d'ordre 1 ou plus) conduisent à un minimum **local**. Par la nature même du procédé de calcul, il est en effet possible que le système se stabilise dans un minimum qui ne soit pas le minimum absolu du système, mais seulement un minimum pour une partie de l'espace des paramètres.

Suivant la position initiale du système dans le paysage énergétique et la topographie de ce dernier, un chemin qui suit les reliefs va emmener le système vers un puits du paysage où il se stabilisera. La notion d'**attracteur** dans un paysage énergétique définit un point de ce dernier qui "attire" à lui les éléments en déplacement. Le **bassin d'attraction** correspond quant à lui à la région du paysage où un système sera capté et attiré par l'attracteur, et se dirigera vers lui au cours de son évolution. Dans notre cas, les puits d'énergie correspondent aux bassins d'attraction d'attracteurs constitués des points de plus faible altitude dans ces puits. Suivant le bassin d'attraction dans lequel se trouvera initialement un système sur le relief, la progression se fera vers l'attracteur correspondant.

Différentes méthodes permettent d'essayer de se sortir de la situation délicate des minima locaux (ceci ne constituera cependant pas un problème pour nous par la suite). Il est tout d'abord possible d'examiner un ensemble de configurations initiales réparties dans l'espace des paramètres. Une seconde possibilité est, lorsque le système s'est stabilisé dans une certaine configuration des paramètres, de tenter des sauts aléatoires, sans tenir compte du relief (autrement dit des barrières énergétiques). Ceci permet de voir si le système, après un saut, peut voir son énergie diminuer par rapport à la valeur du premier minimum. Si tel est le cas, le minimum était local et non absolu, et le système peut recommencer son évolution graduelle et conforme au relief à partir de la nouvelle position atteinte. La méthode emprunte donc à différentes techniques d'ordres différents.

Les figures 6.1 et 6.2 permettent d'illustrer différents types de situation (surfaces dessinées grâce au logiciel Matlab 6.0). Dans le premier cas, un optima absolu existe dans le paysage et son bassin d'attraction (quoique peu marqué) couvre l'ensemble du paysage. Quel que soit le point de départ de l'évolution du système, ce dernier finira par rejoindre au bout d'un certain temps le minimum absolu.

Dans le second cas, deux minima relatifs existent dans le paysage. Suivant la configuration initiale du système et le bassin d'attraction dans lequel il se trouve, il rejoindra l'un des deux minima, comme le montrent les deux évolutions représentées par les trajectoires rouges.

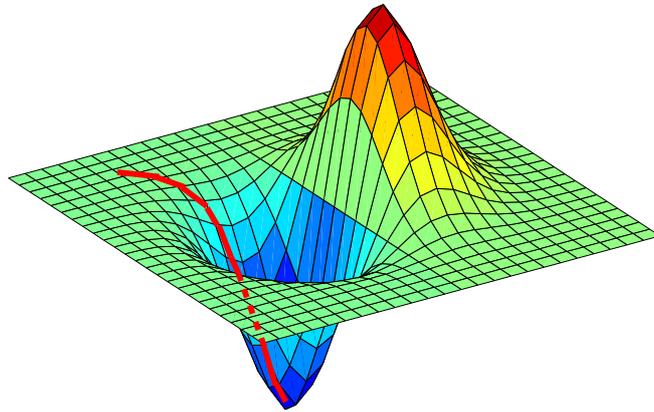


FIG. 6.1 – Paysage énergétique avec un minimum absolu

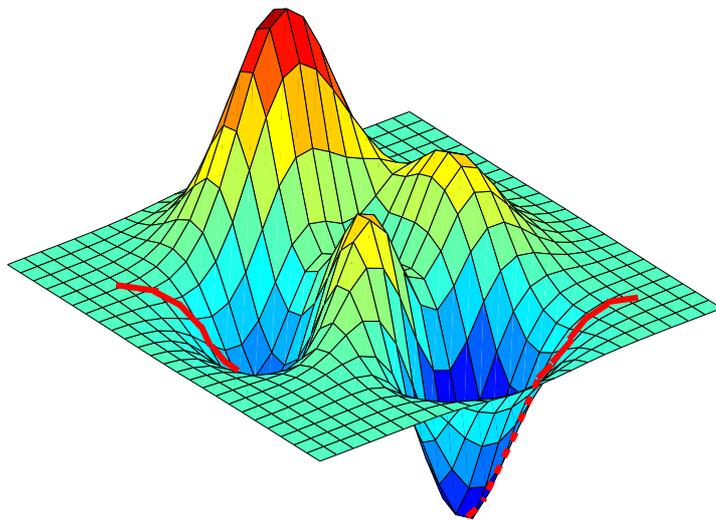


FIG. 6.2 – Paysage énergétique avec deux minima locaux

## Introduction mathématique aux techniques de descente de gradient

Afin d'aborder correctement les paragraphes suivants et le modèle d'évolution linguistique, il nous paraît préférable de commencer par présenter la technique de recherche de minimum de fonction par méthode de *steepest descent* dans un cas général.

Soit une fonction  $f$  de  $n$  paramètres  $(x_1, x_2, \dots, x_n)$ . La méthode de *steepest descent* permet de mettre à jour le minimum local de la fonction "le plus proche" (selon la notion de bassin d'attraction et non de distance euclidienne) d'un jeu de paramètres initial  $(x_1^0, x_2^0, \dots, x_n^0)$ . L'idée est d'envisager une suite de modifications locales des paramètres qui suivent le chemin de plus grande pente du paysage énergétique (déterminé par la fonction  $f$ ) vers le minimum. Dit autrement, chaque modification des paramètres est localement optimale : chaque paramètre varie indépendamment de façon à diminuer la valeur de la fonction  $f$  sur sa dimension.

Une analogie de raisonnement est ici l'algorithme de rétro-propagation du gradient qui permet l'apprentissage des perceptrons multi-couches. Une évolution au cours du temps vers un minimum de la fonction  $f$  se traduit par la condition suivante :

$$\frac{df}{dt} \leq 0$$

La fonction  $f$  est une fonction des paramètres  $(x_1, x_2, \dots, x_n)$ . Puisque nous envisageons une évolution de ces paramètres au cours du temps, ils dépendent également de cette variable, et il est dès lors possible de décomposer la dérivée  $\frac{df}{dt}$  comme suit :

$$\frac{df}{dt} = \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \times \frac{\partial x_i}{\partial t} \right)$$

La première condition devient dès lors :

$$\sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \times \frac{\partial x_i}{\partial t} \right) \leq 0$$

Une des façons de définir l'algorithme de *steepest descent* est de remplacer cette nouvelle condition par l'ensemble des conditions suivantes<sup>40</sup> :

$$\forall i \in [1, \mathbf{n}], \quad \frac{\partial f}{\partial x_i} \times \frac{\partial x_i}{\partial t} \leq 0$$

En définissant un paramètre réel  $\alpha$  positif, il est encore possible de réécrire ces conditions de la façon suivante :

$$\forall i \in [1, \mathbf{n}], \quad \frac{\partial f}{\partial x_i} = -\alpha \times \frac{\partial x_i}{\partial t}$$

<sup>40</sup>Notons ici que la condition peut-être satisfaite de multiples façons, certains produits de la somme pouvant être positifs bien que la somme soit négative. Considérer tous les produits comme négatifs correspond à l'idée de caractère localement optimal des transformations.

L'évolution au cours du temps par *steepest descent* peut dès lors s'écrire :

$$\forall i \in [1, \mathbf{n}], \mathbf{x}_i(t + \partial t) = \mathbf{x}_i(t) - \alpha \times \partial t \times \frac{\partial \mathbf{f}}{\partial x_i}$$

La valeur de  $\alpha$  doit être très faible afin de respecter la localité des changements. Autrement dit, une faible valeur permet d'épouser fidèlement le relief du paysage énergétique. Dans un mode discret (non continu), une valeur trop importante risque de "faire manquer" au système certains optima à cause de sauts de trop forte amplitude. L'optimalité locale n'est alors plus respectée.

Si comme nous l'avons déjà précisé, les méthodes de minimisation d'ordre 1 ou plus cherchent à épouser de façon performante le relief du paysage énergétique pour converger vers un minimum, l'algorithme de *steepest descent* suit en fait le chemin de *plus grande pente*, comme son nom le laisse deviner. Les figures 6.1 et 6.2 illustrent ce type de déplacement sur le paysage. La pente en un point de l'espace des paramètres est donnée par le vecteur dérivé  $\Delta \mathbf{f}$  :

$$\Delta \mathbf{f} = \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial x_1} \\ \frac{\partial \mathbf{f}}{\partial x_2} \\ \vdots \\ \frac{\partial \mathbf{f}}{\partial x_i} \\ \vdots \\ \vdots \end{pmatrix}$$

Les formules données précédemment et ce vecteur permettent de constater comment chaque paramètre  $x_i$  suit bien la pente de sa propre dimension.

### 6.1.2 Objectifs et présupposés théoriques du modèle

Au cours du deuxième chapitre, nous avons abordé les changements linguistiques selon une approche systémique, en mettant l'accent sur les structures linguistiques internes (que ce soit aux niveaux phonologique, lexical, syntaxique... Il s'agit ici de la dimension structurelle.) et sur leur évolution en fonction de contraintes internes (naturelles) et externes (les facteurs sociaux). Nous souhaitons ici développer un modèle simple qui prenne en compte les relations structurelles entre les composants linguistiques et qui soumette leur évolution "structurelle" aux contraintes naturelles et à des phénomènes de type social. Notre but sera entre autres de mesurer l'impact des phénomènes sociaux sur les changements linguistiques.

#### La notion d'item linguistique

La notion d'item linguistique telle que nous l'avons utilisée au cours des chapitres précédents constitue la base de notre modèle.

Comme déjà mentionné au chapitre 2, la notion d'item linguistique est très abstraite, mais elle a l'avantage d'être assez générale pour permettre une approche formelle apte à fournir des résultats quantitatifs et qualitatifs. Elle convient ainsi bien au modèle que nous souhaitons développer.

## Notion de contexte

Un item linguistique prend toujours place dans un **contexte** linguistique particulier. Cette notion est très vaste, et regroupe des niveaux plus ou moins abstraits et généraux : les segments prennent place dans des contextes phonétiques et des morphèmes, les mots du lexique dans des contextes syntaxiques, les phonèmes de la langue dans un système phonologique. . . Un contexte contient ainsi un ou plusieurs items, qui sont éventuellement en compétition avec d'autres items.

Le modèle doit prendre en compte cette notion de contexte. Un contexte peut être occupé par un certain nombre d'items. Dès lors, chacun d'entre eux est représenté dans le contexte selon une **fréquence d'occurrence**. Deux cas sont envisageables : si l'item ne peut être considéré que comme présent ou absent du contexte (par exemple, dans le contexte théorique formé par une voyelle, le trait de nasalité ne peut être que présent ou absent), sa fréquence est soit nulle, soit égale à 1. Si au contraire l'item peut être présent dans le système selon une fréquence plus ou moins importante, celle-ci varie entre 0 et 1.

L'évolution d'un système se traduira par l'évolution des fréquences des différents items en contexte, sous la pression de différentes contraintes que nous décrirons ci-dessous.

Pour un ensemble d'items donné, un contexte linguistique est *a priori* constitué d'autres items linguistiques. Nous ne décrirons cependant pas le contexte en fonction d'une partie des items du système, et postulerons dans notre modèle qu'il existe de façon indépendante des items considérés en son sein. Ceci revient en fait à faire l'hypothèse simplificatrice qu'il existe des stratifications relativement indépendantes le long de la dimension structurelle du système langagier ; même si cette hypothèse n'est pas vérifiée, elle nous permettra de mieux évaluer la situation à l'aide d'un modèle nettement plus simple.

Selon les cas, le contexte pourra être défini comme **exclusif** ou **inclusif**. Dans le cas exclusif, la somme des fréquences des items du contexte devra être égale à 1. Ceci ne sera pas nécessaire dans le cas inclusif. Ces deux possibilités permettent de laisser une certaine liberté sur la définition du contexte linguistique, et ainsi d'envisager différents cas.

Par exemple, si l'on considère le système phonologique d'une langue comme le contexte des différents phonèmes qui le composent, un contexte inclusif permettra de faire co-exister différents éléments de fréquence soit nulle, soit égale à 1. Si l'on considère maintenant le lexique comme une collection de contextes qui représentent les mots de ce lexique, des contextes inclusifs permettront de composer les mots grâce à des segments selon une certaine distribution de fréquences. Ceci permettra d'illustrer par exemple la variabilité phonétique qui peut exister lors de la prononciation de certains mots chez un individu.

## Impact des contraintes naturelles et distribuées sur l'évolution des items linguistiques

L'idée du modèle est de représenter par des outils mathématiques les interactions entre composants des structures linguistiques. Le postulat de base développé dans le chapitre 2 est l'existence d'un certain nombre de contraintes qui pèsent sur ces composants, sur leurs interactions et leurs évolutions. Le terme de **contraintes naturelles et distribuées** est à envisager au sens large, et reflète la façon dont les dimension naturelle et distribuée du système langage viennent peser sur la dimension structurelle de celui-ci. Rappelons ici en particulier que ce sont les contraintes naturelles et distribuées qui viennent définir les structures linguistiques, en sé-

lectionnant les combinaisons d'items cohérentes vis à vis de leur caractéristiques. Parmi ces contraintes, nous pouvons citer :

- des contraintes de production et de perception, liées aux structures physiologiques humaines mises en jeu pour l'utilisation du langage. Il est par exemple plus efficace d'employer des sons faciles à prononcer et à percevoir ;
- des contraintes cognitives, qui correspondent aux aspects cognitifs du langage, comme l'encodage et le décodage des phrases ;
- des contraintes de l'environnement, qui traduisent l'utilisation du langage pour échanger des informations à propos d'objets et de concepts de l'environnement des locuteurs (environnement dont font entre autres partie les autres individus) ;
- des contraintes interactionnelles, qui correspondent à l'idée d'une transmission efficace des informations précédentes grâce au langage.

Afin de rendre compte de ces contraintes naturelles et distribuées, il est possible de définir une notion de **compatibilité** des items, qui caractérisera leur adéquation par rapport aux contraintes (de façon globale) et leur assemblage en structures. Cette compatibilité et les contraintes en jeu sont définies par rapport au contexte dans lequel les items prennent place.

La notion de compatibilité englobe à la fois l'adéquation d'un item aux contraintes hors du cadre défini par un système linguistique particulier, et l'adéquation de ses interactions avec les autres items. S'il est bien sûr possible d'argumenter sur le fait que les items n'apparaissent qu'*en contexte* dans les énoncés linguistiques, il nous paraît pédagogique de distinguer les deux phénomènes précédents, à l'aide des deux dénominations suivantes.

Tout d'abord, chaque item possède une compatibilité propre ou **compatibilité intrinsèque** vis à vis des contraintes naturelles et distribuées. Cette compatibilité est supposée indépendante des autres items du système. Un exemple simple mais abstrait est par exemple d'envisager un segment phonétique particulier, et de mesurer ses facilités de production et de perception hors des contextes de co-articulation (ce qui peut être discutable). La combinaison de ces deux variables permettrait alors d'approcher la compatibilité de ce segment vis à vis d'une partie des contraintes naturelles. Autre exemple, l'examen des fréquences des ordres des mots (SOV, SVO, VSO... voir chapitre 2) pourraient donner des indices sur la compatibilité intrinsèque de ces items linguistiques.

Encore une fois, la possibilité d'isoler un item hors du contexte linguistique des phrases est théorique, mais permet de clarifier la situation. Dans l'exemple précédent, la plus grande rareté de l'ordre OSV [Black, 1999] (p. 19) peut peut-être s'expliquer par des contraintes cognitives qui rendent plus difficile à traiter une phrase dont le sujet et le verbe arrivent après l'objet. Un item linguistique représentant la construction OSV aurait ainsi une moins grande compatibilité intrinsèque que l'item représentant par exemple la construction SOV, mais il est également possible que les plus grandes fréquences de certains ordres soient la conséquence d'interactions systémiques avec d'autres composants structuraux, et qu'elles aient peu à voir avec leur compatibilité intrinsèque.

En plus de leur compatibilité intrinsèque, les items entretiennent des relations entre eux, qui vont elles aussi être soumises aux différentes contraintes naturelles et distribuées (ce qui va conduire à l'apparition de structures). Nous nous proposons de dénommer ces phénomènes par la notion de **compatibilité relationnelle**. *A priori*, estimer le poids des contraintes sur

les interactions entre items requiert l'évaluation de toutes les combinaisons systémiques d'items possibles. Cette description exhaustive est néanmoins difficile à transcrire dans un modèle si le nombre d'items est important, et il est possible de recourir à la simplification suivante : la compatibilité relationnelle peut être définie pour chaque paire d'items, et représenter la compatibilité d'un item vis à vis des contraintes lorsqu'un second item est présent dans le système linguistique. La présence du second item peut entraîner une bonne ou une mauvaise compatibilité du premier item. Nous pouvons citer ici l'exemple de la relation entre ordre des mots et position de la tête que nous avons introduite au chapitre 2.

Cette définition simplifiée de la compatibilité relationnelle repose sur l'hypothèse qu'il est possible de considérer la compatibilité d'un item en co-occurrence avec un système d'items grâce à une décomposition de la compatibilité relationnelle globale en compatibilités deux à deux avec les différents items du système linguistique.

Cette dernière notion n'est *a priori* pas symétrique mais ordonnée : nous distinguerons ainsi la compatibilité relationnelle d'un item  $i$  en présence d'un item  $j$  de la compatibilité relationnelle de l'item  $j$  en présence de l'item  $i$  dans le système. L'idée sous-jacente est que si un item linguistique favorise l'apparition d'un autre item, la réciproque n'est pas forcément vraie (la notion de compatibilité ne recouvre cependant pas la notion de facilité d'émergence, comme nous le verrons par la suite). La relation porte donc en fait non pas sur une paire d'éléments non-ordonnés, mais sur chaque item en fonction de la présence ou non d'autres items dans un système. Une des conséquences de cette proposition est la possible existence d'universaux *implicationnels*, comme ceux rencontrés dans les langues du monde.

Si la compatibilité relationnelle est définie pour chaque paire *ordonnée* d'items, nous pouvons enfin utiliser les termes de **compatibilité en situation** ou **compatibilité systémique** d'un item pour désigner la compatibilité globale de celui-ci dans un système particulier, déterminée par sa compatibilité intrinsèque et la présence de l'ensemble des autres items qui forment le système. Le terme de **cohérence systémique** définira quant à lui l'adéquation aux contraintes, non pas d'un item particulier en interaction avec d'autres items, mais celle d'un ensemble d'items, c'est à dire d'un (sous-)système linguistique. Cette cohérence pourra être construite à partir des compatibilités intrinsèques et relationnelles des éléments qui composent le système.

Une compatibilité pourra être négative, positive ou nulle. Dans le premier cas, elle traduira le fait qu'un item va à l'encontre des forces du système (plus ou moins fortement), dans le second cas, que sa présence va dans le sens des contraintes qui pèsent sur le langage (plus ou moins fortement), et dans le troisième cas, que sa présence est neutre vis à vis du jeu de contraintes.

Deux approches peuvent être envisagées pour mettre en jeu la compatibilité des items. Une première approche très abstraite, dont le but est plus d'investir les phénomènes dynamiques de l'évolution d'un système sous contraintes, présuppose une fonction d'énergie donnée *a priori* pour le système linguistique. Les valeurs de cette fonction pour l'ensemble des configurations du système linguistique définissent la cohérence systémique de celui-ci dans son ensemble de façon immédiate, sans décomposition selon les items. Ce cas est le plus général possible, puisque tout repose sur une unique fonction, qui capture l'ensemble des contraintes qui pèsent sur les items et leurs interactions. Cependant, il est bien sûr presque impossible de définir cette fonction pour un objet aussi complexe que le langage. Nous référerons à cette approche grâce à l'expression **approche holistique**, dans le sens où elle ne décompose pas la cohérence du système selon les compatibilités des éléments qui le composent.

Une seconde approche, moins générale et donc moins puissante, mais peut-être plus utile pour approcher des phénomènes linguistiques réels, consiste à faire apparaître de façon explicite les différentes compatibilités que nous avons définies plus haut. Nous appellerons cette approche **approche combinatoire**.

Dans ce chapitre, nous nous concentrerons sur l'approche holistique et sur les aspects dynamiques de l'évolution des systèmes linguistiques, qui peuvent être étudiés à l'aide d'un modèle abstrait. Nous présenterons toutefois en annexe les principes et les outils mathématiques pour implémenter l'approche combinatoire, que nous souhaitons mettre à profit dans le futur pour ancrer notre modèle dans la réalité (voir les conclusions du chapitre 7).

### Situation mono- ou pluri-contextuelle

Nous avons introduit plus haut la notion de contexte pour l'existence d'items dans un système linguistique. Nous avons également écrit que les différentes notions de compatibilités et de contraintes étaient définies par rapport à un contexte particulier.

Il est tout à fait possible d'envisager la co-existence de plusieurs contextes de type similaire ou non dans le système : par exemple, les mots du lexique représentent des contextes de type similaire pour les segments, de même que différentes structures syntaxiques peuvent accueillir les mots d'une langue particulière.

Parallèlement aux contraintes naturelles et distribuées prises en compte par les différentes compatibilités des items, des **contraintes contextuelles** doivent être ajoutées au modèle lorsque plusieurs contextes co-existent.

Il est important de bien voir ici que ces contraintes contextuelles font elles aussi partie des contraintes naturelles et distribuées qui pèsent sur les systèmes linguistiques, mais qu'elles interviennent à un niveau différent des contraintes précédentes, qui étaient elles spécifiques à un contexte donné. Pour reprendre l'exemple des segments dans les mots du lexique, des contraintes naturelles et distribuées pèsent sur les segments à l'intérieur des mots (qui représentent des contextes de nature identique), tandis que d'autres pèsent sur les mots eux-mêmes.

La distinction entre les deux dénominations utilisées n'est donc pas très pertinente, puisque les contraintes contextuelles sont elles aussi naturelles ou distribuées, mais ces deux expressions serviront à simplifier notre discours, ainsi qu'à envisager simplement les cas où un seul contexte est en jeu.

Si des items peuvent apparaître dans des contextes de nature différente, alors pour chaque type de contexte, il sera nécessaire de définir les contraintes qui pèsent sur l'item *dans* ce contexte. Par soucis de simplicité, nous n'envisagerons par la suite que des contextes de nature similaire, c'est à dire dans lesquels les contraintes naturelles et distribuées qui pèsent sur les items sont identiques.

Notre proposition est que les différents contextes entretiennent entre eux des relations de répulsion ou d'attraction, de la même façon que les individus entre eux (rappelons que nous souhaitons souligner les similarités entre dimensions distribuée et interne du langage au cours de notre travail). Un exemple assez abstrait est le suivant : les mots ou plutôt les différentes syllabes ou plus généralement "sous-parties" qui composent les mots du lexique forment un ensemble de contextes pour les sons de la langue. Il existe une certaine pression pour que les sons dans des

contextes identiques soient similaires, afin de simplifier la mémorisation du lexique, mais également des contraintes pour éviter des phénomènes comme une homophonie trop importante. Le phénomène de diffusion lexicale, où des changements se répandent de façon graduelle dans le lexique d’une langue, traduit selon nous une pression due à une certaine contrainte de similarité des sons dans des contextes similaires (par exemple les consonnes finales dans les dialectes chinois), liée à des phénomènes cognitifs sous-jacents.

### Probabilité d’émergence et stabilité des items

Dans le modèle, aucune assertion directe n’est formulée quant à la stabilité et à la “facilité d’apparition” des items. De telles notions ont été discutées par Greenberg en lien avec l’existence d’universaux linguistiques et la fréquence de certains éléments typologiques dans les langues du monde [Greenberg, 1978]. Nous avons détaillé ces propositions au chapitre 2.

Dans le modèle, les cas de figures proposés par Greenberg doivent logiquement émerger comme “produits dérivés” des différentes interactions structurelles et des compatibilités intrinsèques des items, et il sera possible de déterminer sous quelles conditions les différents scénarios sont réalisés. Nous avons déjà souligné plus haut que la compatibilité d’un item vis à vis d’un système était en lien avec les possibilités d’apparition et de persistance de cet item dans le système, ou au contraire avec celle de son éviction du système, mais la relation n’est pas forcément directe et linéaire. Plus généralement, les évolutions du système résultent en partie des compatibilités intrinsèques et relationnelles des items linguistiques.

### Prise en compte des facteurs sociolinguistiques

**Système linguistique d’une communauté d’individus.** Deux approches sont envisageables pour étudier l’évolution d’un “système linguistique” : soit on construit un système abstrait représentant une langue, sans se soucier des locuteurs et des différences inter-individuelles, soit on considère l’ensemble des locuteurs et leurs proximités linguistiques et sociales. Ceci revient en fait à proposer un choix pour la définition d’un système linguistique : soit celui-ci représente une langue de façon abstraite, soit chaque individu possède son propre système linguistique (un idiolecte), plus ou moins proche de ceux des autres individus. Nous pouvons rebondir ici sur le concept de langage communal introduit au chapitre 2, et envisager de construire une représentation moyenne des systèmes linguistiques des individus d’une communauté.

Dans le premier cas, il n’est pas nécessaire d’envisager les constructions sociales de façon directe, mais elles peuvent néanmoins être incluses de façon indirecte dans le modèle. Les prochains paragraphes concernent quant à eux le second cas.

**Modélisation des structures sociales.** Afin d’ancrer notre modèle dans la réalité des interactions sociales qui existent entre les individus, il nous paraît utile et nécessaire de nous appuyer sur des modèles théoriques simples de celles-ci. Cette approche est rare en modélisation de l’évolution du langage, comme nous l’avons vu au chapitre 3, mais des modèles ont été proposés par les sociolinguistes, comme celui des réseaux de Milroy que nous avons brièvement présenté au chapitre 2.

La représentation de la dimension distribuée du système langage passera donc par la modélisation de réseaux sociaux entre les individus d’une communauté linguistique. Les liens pourront être plus ou moins forts, et caractériser une attraction ou une répulsion. Ils ne seront pas nécessairement symétriques, ce qui traduira le fait que la façon dont un individu considère une

seconde personne peut ne pas être réciproque. Des situations comme l'existence d'individus très influents mais qui ne prêtent que peu d'attention aux autres sont ainsi très faciles à mettre en place. Le point de vue suivant de Labov résume la "philosophie" du modèle :

*"Like linguists, sociologists are prone to explain social behavior by inferring a series of covert attitudes for which the main evidence is the behavior itself. Anspach in The Why of Fashion (1967) argues that "the initiating spark is the need of people to be like others and yet to be distinct from others" (pp. 5-6). A theory that embraces the need to be like others and the need to be distinct from others is capable of explaining any combination of events."* [Labov, 2001] (p. 361)

L'intérêt de ces liens réside bien sûr en ce qui nous concerne dans leurs conséquences au niveau linguistique. Nous reprenons partiellement une des caractéristiques du principe de densité de Bloomfield, édicté en 1933, et que Labov rapporte de la façon suivante :

*"... from Bloomfield's principle of density, which argues that each communicative act is accompanied by a slight degree of convergence of linguistic systems of speaker and interlocutor."* [Labov, 2002].

Le but de notre modèle n'est pas de détailler le contenu des interactions linguistiques, mais de regarder comment les différents liens sociaux et les contraintes naturelles et distribuées viennent peser sur les changements linguistiques. L'adoption de principes généraux et plus abstraits comme celui de Bloomfield permet de regarder l'évolution du système à un niveau global, sans rentrer dans les extrêmes complications d'une modélisation des interactions sous forme syntaxique, lexicale...

Nous étendons légèrement le principe de Bloomfield tel qu'il est rapporté par Labov en supposant que la convergence lors des actes communicatifs se produit lorsque le lien entre les individus est attracteur (les deux individus se reconnaissent comme proches socialement), et qu'au contraire une légère divergence se produira lors d'une interaction entre deux locuteurs unis par un lien de valeur négative (les deux individus s'opposent partiellement sur le plan social). En outre, comme cela était le cas dans la *Social Impact Theory* mise à profit par Nettle, nous postulons que l'intensité de la convergence ou de la divergence linguistique est reliée de façon *non linéaire* à l'intensité du lien entre les deux individus : les convergences ou divergences linguistiques croissent de moins en moins vite lorsque l'attraction ou la répulsion sociale augmente.

La propagation des changements linguistiques dans une communauté va ainsi reposer sur le schéma des liens entre individus, et nous pouvons parler en un certain sens de **contraintes sociales**. Comme nous l'avons déjà souligné, de très nombreuses situations peuvent être envisagées, et nous espérons pouvoir les examiner plus précisément (en définissant des typologies de schémas sociaux) dans un futur proche. Nous nous limiterons cependant dans ce chapitre à des cas simples, afin de ne pas nous détourner des résultats les plus fondamentaux que nous souhaitons mettre en exergue.

### Distribution des changements vis à vis des contraintes internes et externes

Deux opinions s'opposent souvent dans la littérature linguistique. D'un côté, certains articles ou travaux insistent sur certains changements en démontrant comment (ceci est notre interprétation vis à vis de notre cadre théorique) ils peuvent entraîner une meilleure adéquation du

système linguistique aux contraintes. D'un autre côté, les études sociolinguistiques ou sur les contacts linguistiques insistent sur le fait que des changements de tout type peuvent se produire, et qui peuvent transformer des composantes structurelles *a priori* très centrales des langues : ordre des mots, systèmes tonaux. . .

Cette opposition conduit donc à première vue à devoir choisir entre une absence de contraintes dans la sélection des transformations qui pèseront sur les langues, ou une existence de changements qui vont toujours dans le sens d'une meilleure adéquation des langues vis à vis des contraintes internes (naturelles et distribuées).

Nous proposons d'opter pour une **distribution probabiliste** des changements qui touchent les langues, vis à vis des multiples contraintes, externes (sociales) ou internes (naturelles, distribuées ou contextuelles), qui pèsent sur eux. Le but de cette distribution est simplement de traduire l'idée que des changements plus en adéquation avec l'ensemble des contraintes auront tendance à se produire plus souvent que des changements qui déstabilisent le système. Nous calculerons donc concrètement dans nos simulations une distribution des changements possibles à partir des contraintes naturelles et distribuées, contextuelles et sociales qui pèsent sur l'évolution d'un système linguistique. Cette distribution représentera la façon dont chaque changement potentiel répond à l'ensemble des contraintes. Une projection sera alors effectuée entre la pertinence des changements potentiels vis à vis des contraintes et la probabilité qu'ils se réalisent dans le système linguistique.

La raison intuitive de ce choix est la suivante : à chacune de leur utilisation, les formes linguistiques se trouvent confrontées au jeu de l'ensemble des contraintes, et sont donc évaluées vis à vis de lui. Même si des processus contingents sont toujours à l'œuvre, l'apparition et la persistance d'un changement sont donc le résultat d'une évaluation permanente vis à vis des contraintes. Selon l'orientation et la force respectives des contraintes naturelles et distribuées, contextuelles ou sociales, certains changements pourront apparaître plus ou moins facilement, des transformations persister plus ou moins longtemps, et se répandre ou non dans une communauté de locuteurs.

Par exemple, des changements introduisant une perte de stabilité du système vis à vis des contraintes internes, et donc une plus grande difficulté d'utilisation, pourront apparaître et persister pour caractériser la position sociale de certains locuteurs, mais auront des difficultés à s'étendre au-delà du cadre de la sous-communauté où les processus d'identification sociale les rendent possibles et persistants.

Notons ici que l'idée d'une distribution des changements peut s'appliquer au système linguistique d'un locuteur comme au système linguistique d'une langue abstraite. Par le biais du phénomène d'implémentation des changements dans une communauté, on peut espérer une certaine similarité entre les deux niveaux considérés.

### Différents scénarios d'évolution

Si l'on se réfère à l'opposition partielle soulevée dans les derniers paragraphes précédents, ainsi qu'à notre proposition de distribution des changements potentiels, nous pouvons envisager différents scénarios d'évolution linguistique. Notre but sera bien sûr d'examiner les dynamiques d'évolution résultantes, et de les contraster avec la situation des langues réelles.

Un premier point préliminaire à souligner est qu'hormis peut-être dans des conditions sociales

et linguistiques très extrêmes, les évolutions du système sont le plus souvent **locales**. Ce caractère est d'importance, et peut se justifier de la façon suivante : hormis de très rares cas, les locuteurs n'ont pas conscience de la façon dont les langues évoluent, ni de comment les contraintes pèsent sur les changements linguistiques, même s'ils peuvent détecter des variations dans les différents parlers qui les entourent ou dans le leur propre. Si des changements se produisent dans un système linguistique, qu'ils stabilisent ou déstabilisent celui-ci (en augmentant ou diminuant l'adéquation aux contraintes), ils porteront sur un faible nombre d'items, en réaction à des caractéristiques locales du paysage énergétique et/ou pour manifester des phénomènes sociaux. Les locuteurs n'ont en effet pas l'intuition que toute une cascade de changements pourrait éventuellement conduire à une amélioration ou à une détérioration tangible de leur système. Notons toutefois qu'une telle réflexion n'est pas impossible, et que des langues construites de façon théorique comme l'espéranto avaient pour but d'être aisément apprises par un grand nombre de locuteurs de langues natives différentes (même si le problème d'une adéquation à un jeu de contraintes n'était pas posé en ces termes, le principe est assez proche).

Pour énoncer la situation en d'autres termes, la langue et les locuteurs n'ont aucune idée des puits d'énergie du paysage énergétique du langage, et *a fortiori* aucune idée de l'existence et de la localisation de minima absolus ou très importants. Le schéma d'évolution possible repose sur un **mécanisme** de mouvement dans l'espace des possibles linguistiques, en utilisant (de façon inconsciente) les informations sur la *topologie locale* du paysage énergétique à l'emplacement du système : énergie, valeurs locales des dérivés de la surface qui renseignent sur les directions et les inclinaisons des pentes du paysage énergétique. Ceci correspond en fait à l'évaluation permanente des formes linguistiques, et c'est par une itération du mécanisme d'évolution locale que se construit l'évolution au cours du temps.

Il est tout d'abord possible de s'intéresser à des évolutions du système qui cherchent à maximiser de façon continue et **localement optimale** la cohérence systémique du système linguistique vis à vis des contraintes. Dans une vision énergétique de la situation, le système se dirige vers l'attracteur, situé au fond d'un puits d'énergie, dont le bassin d'attraction contient sa configuration initiale. En outre, le déplacement se fait d'autant plus rapidement que la pente menant à l'attracteur est importante.

Si l'on se réfère au cadre mathématique défini au début du chapitre, il est possible de mathématiser la situation précédente à l'aide de dérivées partielles, et donc d'obtenir une description **continue** du mécanisme. Néanmoins, cette approche est idéale, et dans les techniques d'optimisation réelles, comme vraisemblablement également pour les phénomènes linguistiques, l'aspect continu est remplacé par des sauts discrets de faible portée. La faible distance parcourue à chaque saut permet de ne pas briser le caractère (idéal et) local des évolutions. Les propriétés de convergence vers des états stables des méthodes d'optimisation détaillées plus haut sont ainsi conservées.

Nous pouvons recourir au terme de **mutations** pour désigner ces petites transformations discrètes du système linguistique. Une mutation fait passer le système d'un état **A** à un état **B** voisin, et peut donc être définie par le vecteur reliant ces deux états. Dans le cas d'items ne pouvant être que présents ou absents, les mutations font passer leur fréquence d'une valeur à une autre (de 0 à 1, ou de 1 à 0). Dans le cas de fréquences pouvant varier continûment entre 0 et 1, les mutations peuvent être plus ou moins importantes, allant de quelques pourcents à des changements plus radicaux. Dans ces derniers cas, l'aspect local de l'évolution vis à vis d'un item pourra être remis en cause, mais globalement, l'évolution du système restera locale vis à vis du vaste ensemble d'items linguistiques qui composent celui-ci.

Dans le premier scénario précédent, les vecteurs des mutations sont toujours proportionnels aux vecteurs de plus grande pente du système (méthode de *steepest descent*). Toutefois, d’après ce que nous venons de dire plus haut, nous envisagerons en fait des mutations partiellement stochastiques au niveau temporel : si le résultat au bout d’un temps infini d’évolution sera le même que dans le premier cas continu, le décours temporel des changements sera lui stochastique.

Deux autres scénarios peuvent être construits en faisant varier les directions des vecteurs de mutation.

Un second scénario consiste à faire évoluer le système selon des mutations aléatoires, sans se soucier des pentes du paysage énergétique. Ceci correspond à une évolution sans contraintes du système. Le scénario d’évolution sera alors à la fois stochastique au niveau temporel et au niveau des variations des paramètres du système.

Enfin, un troisième scénario correspondra à notre proposition de **distribution** des changements. Les directions des mutations se feront donc selon la distribution probabiliste utilisée : la majeure partie d’entre elles aura donc tendance à renforcer l’adéquation du système vis à vis des contraintes, mais certaines plus rares iront au contraire dans la direction opposée. Les conséquences dynamiques de cette évolution seront particulièrement intéressantes à observer.

### 6.1.3 Description mathématique du modèle

Nous ne présentons ici que la mathématisation de l’approche holistique, plus aisée que celle de l’approche combinatoire décrite en annexe.

#### Compatibilité et énergie d’un système linguistique

Notre hypothèse de départ est l’existence d’un ensemble d’items linguistiques fini, composé de  $n$  éléments, dans lequel une langue va pouvoir “piocher” un certain nombre d’éléments afin d’être apte à véhiculer de l’information entre ces locuteurs (cette définition qui fait de la langue une entité presque douée de volonté n’est bien sûr qu’une figure de style).

Comme nous l’avons dit plus haut, nous supposons dans l’approche holistique que nous disposons d’une fonction d’énergie qui donne pour chaque configuration d’un ensemble d’items une valeur numérique qui correspond à l’adéquation du système vis à vis du jeu de contraintes naturelles et distribuées. Dans les expériences développées plus bas, nous ne considérons que des items dont la fréquence peut varier de façon continue dans le ou les contextes.

Nous définissons l’énergie d’un système linguistique comme l’inverse de son adéquation vis à vis des contraintes, ce qui permet de parler de minimisation comme dans la plupart des cas d’étude d’autres domaines (même si minimisation et maximisation sont deux processus tout à fait symétriques).

Rappelons ici (voir chapitre 2) que l’espace énergétique engendré par la fonction d’énergie va être bâti sur l’espace des états possibles du système linguistique. Par exemple, si l’on envisage deux items dont les fréquences peuvent varier entre 0 et 1, ce qui sera le cas dans toutes les expériences suivantes, alors l’espace des états possibles sera composé de tous les doublets possibles (a,b) choisis dans l’intervalle  $[0, 1] \times [0, 1]$ . L’espace énergétique sera lui composé en ajoutant une troisième dimension qui associe une valeur d’énergie à chaque couple de fréquences.

### Représentation d'un contexte et d'un ensemble de contextes

Il nous faut maintenant définir une représentation pour la notion de contexte que nous avons introduite plus haut. La modélisation la plus simple consiste en un vecteur de  $\mathbf{n}$  éléments caractérisant les fréquences d'occurrence des différents items linguistiques dans ce contexte. Les fréquences varient entre  $\mathbf{0}$  et  $\mathbf{1}$ , une valeur nulle traduit l'absence de l'item linguistique dans le contexte, alors qu'une valeur de  $\mathbf{1}$  traduit la présence permanente de cet item. Selon la caractéristique inclusive ou exclusive du contexte, la somme des fréquences dans un contexte devra ou non être égale à 1.

Dans le prolongement de la notion de contexte unique, un système linguistique pourra être défini comme une collection de contextes. Là où un simple vecteur détermine un contexte, un système linguistique composé de  $\mathbf{q}$  contextes sera dès lors représenté par une matrice de dimensions  $\mathbf{n} \times \mathbf{q}$ , où chaque colonne représentera un contexte, et chaque ligne  $\mathbf{i}$  les fréquences de l'item  $\mathbf{i}$  dans les différents contextes.

$$\mathbf{V} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_i \\ \vdots \\ \vdots \end{pmatrix}$$

$\mathbf{V}[\mathbf{i}] = \mathbf{f}_i =$  fréquence de l'item  $\mathbf{i}$  dans le contexte  $\mathbf{V}$

$$\mathbf{S} = \begin{pmatrix} f_{1,1} & f_{1,2} & f_{1,3} & \dots & \dots & \dots \\ f_{2,1} & f_{2,2} & f_{2,3} & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & f_{i,j} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$\mathbf{S}[\mathbf{i}, \mathbf{j}] = \mathbf{f}_{i,j} =$  fréquence de l'item  $\mathbf{i}$  dans le contexte  $\mathbf{j}$

Les contextes peuvent être indépendants ou non. Pour chaque couple de contextes, une valeur numérique réelle définit le **couplage** entre ces deux contextes. Une valeur positive définit un phénomène d'**attraction** entre les deux contextes, qui auront tendance à rapprocher les fréquences de leurs items linguistiques au cours du temps (phénomène d'**homogénéisation**). Si la valeur est négative, une répulsion se produira entre les deux contextes qui auront tendance à faire diverger les fréquences de leurs items respectifs. Nous supposons ici que la relation de couplage entre deux contextes est symétrique.

Il est ainsi possible de définir les relations entre un ensemble de contextes grâce à une *matrice de relations contextuelles*  $\mathbf{RC}$ . Cette matrice de dimensions  $\mathbf{q} \times \mathbf{q}$  sera triangulaire supérieure ou inférieure et à diagonale nulle, en conséquence de la symétrie de la relation de couplage.

## Représentation des liens sociaux entre individus

Pour les simulations où nous envisagerons une population de locuteurs et non une “langue” abstraite, il sera nécessaire de définir les liens sociaux entre les individus.

L’implémentation informatique est la même que pour un ensemble de contextes. Nous définissons la relation entre deux individus  $\mathbf{I}_1$  et  $\mathbf{I}_2$  grâce à deux valeurs numériques réelles qui représentent l’influence de  $\mathbf{I}_1$  sur  $\mathbf{I}_2$  et celle de  $\mathbf{I}_2$  sur  $\mathbf{I}_1$ . Une valeur positive traduit une attraction, qui correspond en fait à une proximité sociale, et une valeur négative une répulsion. La valeur absolue traduit la force de l’attraction ou de la répulsion.

Pour une population de  $\mathbf{m}$  individus, La *matrice de relations individuelles*  $\mathbf{RI}$ , de dimensions  $\mathbf{m} \times \mathbf{m}$ , regroupe l’ensemble des relations précédentes : la valeur de la cellule  $\mathbf{RI}[\mathbf{i}, \mathbf{j}]$  correspond à l’influence de l’individu  $\mathbf{j}$  sur l’individu  $\mathbf{i}$ . Contrairement à la matrice de relations contextuelles, la matrice de relations individuelles n’est pas triangulaire. Toutefois, sa diagonale est également nulle<sup>41</sup>.

## Evolution de la structure d’un système linguistique : sélection des changements

Nous atteignons ici le cœur de notre modèle, puisqu’il s’agit de traduire dans l’évolution du système linguistique à la fois l’influence des contraintes naturelles et distribuées et de celles du milieu social.

Dans le cadre de l’approche holistique et de sa représentation grâce à des paysages énergétiques, l’évolution d’un système linguistique va correspondre à une trajectoire à la surface du paysage. Notons toutefois que la visualisation d’une trajectoire à une date  $t$  ne fournit aucune indication sur le décours de l’évolution au cours du temps, et que d’autres indicateurs doivent être mis à profit pour bien analyser l’évolution d’un système.

Chaque position sur le paysage énergétique correspond à une configuration particulière du système. Le paysage d’un système de  $\mathbf{n}$  items est de dimension  $(\mathbf{n} + 1)$  (un espace des paramètres de dimension  $\mathbf{n}$  et une dimension supplémentaire pour la valeur de l’énergie). Il est impossible de se représenter visuellement un espace à plus de trois dimensions, mais l’observation des évolutions en dimension 2 ou 3 nous permettra une extrapolation conceptuelle aux dimensions supérieures, ceci à cause de l’indépendance des dynamiques d’évolution vis à vis du nombre d’items du système linguistique.

Afin d’expliquer de façon claire la façon dont les changements linguistiques vont s’opérer, nous allons commencer par aborder le problème d’une façon générale, avant de montrer comment les contraintes naturelles et distribuées, contextuelles et sociales s’y intègrent.

Un changement linguistique se traduit par une variation de la fréquence d’un item dans un contexte linguistique. De nombreuses variations (en fait une infinité dans le cas continu), plus ou moins importantes, et dans un sens ou dans l’autre (vers une fréquence plus faible ou plus forte), peuvent être envisagées. Supposons que la qualité de ces variations soient représentée par une distribution continue ou discrète (le cas échéant) de probabilités  $\mathbf{d}$  : plus une variation d’intensité  $i$  et de sens  $s$  améliorera la “qualité” du système, donc sera de bonne qualité, plus la

<sup>41</sup>On peut également envisager une diagonale non nulle pour doter chaque individu d’une certaine inertie au changement.

valeur de la distribution sera importante pour cette variation. Nous souhaitons que la probabilité qu'elle soit choisie lors d'une évolution du système soit alors plus importante.

Concrètement, étant donné une distribution de qualité pour les changements possibles d'un système, nous postulons que la probabilité d'apparition effective d'un changement est proportionnelle à sa qualité. A chaque pas de temps, un jeu de tirages aléatoires permettra de déterminer si oui ou non le changement choisi se produit effectivement dans le système. Le caractère contingent des changements se traduit ici par l'utilisation de ce tirage aléatoire.

Grâce à cette méthode, nous obtenons à la fois une distribution des changements et un déroulement temporel de leur apparition influencés par leur qualité.

Nous pouvons maintenant rentrer plus en détail et décomposer cette notion de qualité. Il s'agit en fait de l'adéquation à l'ensemble des contraintes externes (sociales) et internes (naturelles, distribuées et contextuelles) qui pèsent sur le système linguistique. Chacun de ces trois ensembles de contraintes va en fait définir sa propre distribution de qualité pour les changements possibles. La distribution finale utilisée pour déterminer l'apparition des changements devra donc être la composition des sous-distributions de chaque ensemble de contraintes.

Les paragraphes suivants expliquent comment calculer la distribution de chaque sous-ensemble.

### Distribution liée au jeu de contraintes naturelles et distribuées

En chaque point du paysage énergétique des contraintes naturelles et distribuées, l'étude de la topologie locale du paysage permet de connaître comment le système peut évoluer pour obtenir une meilleure adéquation aux contraintes, ou au contraire une moins bonne. Le gradient nous donne en effet les pentes du paysage pour les différentes dimensions de l'espace des paramètres.

Nous nous intéressons aux évolutions de chaque item séparément, mais la topologie locale reflète néanmoins la qualité des possibles évolutions de chaque item vis à vis à des autres.

Pour chaque item du système, nous connaissons donc la pente du paysage. Une pente très importante caractérise un système très déséquilibré pour l'item considéré. Au contraire, une pente très faible indique un état d'équilibre. Dans ce dernier cas, peu de pression existe sur les possibles évolutions de la fréquence de l'item. Au contraire, dans le premier cas, aller dans le sens de la pente permettra de mieux répondre aux contraintes de façon significative, et aller en sens inverse augmentera d'autant l'inadéquation du système.

A partir de ce constat, nous bâtissons une distribution **gaussienne** de qualité des changements possibles de moyenne proportionnelle à la valeur de la pente. L'écart-type est lui fixé de façon externe par l'utilisateur. Comment justifier ce choix ? Tout d'abord, le choix d'une distribution gaussienne peut se justifier par le fait que les transformations du système linguistique, de nature locale, répondent souvent très finement aux déséquilibres structuraux, mais également plus rarement de façon excessive. La forme de la Gaussienne permet l'occurrence rare de ces mouvements excessifs.

En prolongeant ce raisonnement, le choix de la moyenne de la gaussienne traduit le fait que dans une situation de déséquilibre, la réponse du système sera *le plus souvent* (mais pas toujours) un changement allant dans le sens d'une meilleure adéquation. Plus le système sera en déséquilibre, plus probables seront les transformations tentant d'améliorer l'adéquation aux contraintes, et plus rares seront celles brisant encore un peu plus cette adéquation.

Nous avons envisagé plus haut trois scénarios possibles de réponse aux contraintes naturelles et distribuées du système. Deux de ceux-ci correspondent en fait à des cas extrêmes dans le choix de l'écart-type de la distribution gaussienne. Le premier scénario, qui consiste à toujours aller dans le sens de la pente, et à une vitesse proportionnelle à l'importance de cette pente, consiste

simplement à choisir un écart-type nul pour la distribution gaussienne. Au contraire, le second scénario, qui postule une indépendance vis à vis des contraintes, correspond à un écart-type infini : toutes les situations sont d'égale qualité.

Le troisième scénario se trouve à mi-chemin entre les deux premiers, avec une valeur intermédiaire de l'écart-type.

### Distribution liée au jeu de contraintes sociales

La façon dont nous allons déterminer la distribution de qualité des changements potentiels vis à vis des contraintes sociales va être très similaire à la méthode précédente. Le raisonnement ne concerne qu'un seul contexte pour plus de simplicité, mais s'étend facilement à plusieurs.

Pour chaque agent  $i$  et chaque item  $k$ , il est possible d'inspecter la situation de l'item chez les autres agents en tenant compte des liens entre ceux-ci et le premier l'agent. Il est possible de définir et calculer une **déviatiion sociale optimale** pour l'agent  $A_i$  de l'item selon la formule suivante :

$$\mathbf{d}_s(k, A_i) = \sum_{j=1, j \neq i}^m d_s^l(k, A_i, A_j)$$

Le calcul de la fonction  $d_s^l(k, A_i, A_j)$  va être différent selon la valeur du lien entre les deux agents  $A_i$  et  $A_j$ , c'est à dire selon la valeur de  $RI_{i,j}$ .

Dans le cas où le lien est nul, la fonction vaut 0. Il n'y pas d'influence de l'agent  $A_j$  sur l'agent  $A_i$ .

Dans le cas où  $RI_{i,j}$  est positif, nous appliquons les formules suivantes, où  $signe()$  est une fonction qui rend +1 pour une valeur positive, et -1 pour une valeur négative :

$$c = th(RI_{i,j}) \times (f_k(A_j) - f_k(A_i))$$

$$\mathbf{d}_s(k, A_i, A_j) = signe(c) \times 0.05 \times \left( 1 - e^{-\frac{c^2}{2 \times 0.3}} \right)$$

Dans le cas maintenant où  $RI_{i,j}$  est négatif, deux nouvelles situations doivent être distinguées selon que les fréquences de l'item  $k$  chez les deux agents sont égales ou non. Dans le cas d'une égalité, nous procédons selon un tirage aléatoire conduisant en moyenne à une probabilité égale de choisir l'une ou l'autre des deux formules suivantes :

$$\mathbf{d}_s(k, A_i, A_j) = 0.05 \times \left( 1 - e^{-\frac{th(RI_{i,j})^2}{2 \times 0.3}} \right)$$

$$\mathbf{d}_s(k, A_i, A_j) = -0.05 \times \left( 1 - e^{-\frac{th(RI_{i,j})^2}{2 \times 0.3}} \right)$$

Enfin, dans le cas d'une inégalité, nous appliquons les formules suivantes :

$$c = th(RI_{i,j}) \times (1 - |f_k(A_j) - f_k(A_i)|)$$

$$\mathbf{d}_s(k, A_i, A_j) = \text{signe}(f_k(A_i) - f_k(A_j)) \times 0.05 \times \left(1 - e^{-\frac{e^2}{2 \times 0.3}}\right)$$

**th** est une fonction tangente hyperbolique définie comme suit :

$$\mathbf{th}(x) = \frac{1 - e^{-x}}{1 + e^x}$$

Dans le cas d'un lien positif entre deux agents, la déviation tendra à rapprocher les agents, tandis que dans le cas d'un lien négatif, elles tendra à les repousser. Dans ce dernier cas, si les fréquences de deux individus sont égales, le sens de la déviation fréquentielle sera déterminé au hasard. La déviation est d'autant plus forte que le lien est fort en valeur absolue, mais avec une atténuation progressive de l'effet lorsque les valeurs de ce dernier augmentent. La somme des déviations donne une déviation moyenne qui traduit l'ensemble des interactions de l'individu avec les autres.

Grâce à la déviation optimale précédente, il est possible de définir une distribution de qualité des changements fréquentiels possibles. Le plus simple est encore une fois de considérer une distribution gaussienne centrée sur la déviation optimale, et avec un écart-type fixé par l'utilisateur. En moyenne, les individus auront ainsi tendance à aller dans le sens des contraintes sociales, mais sans exclure des événements plus rares allant en sens contraire, et ce d'autant plus que la variance sociale sera importante.

Rappelons ici que la variance est en fait le carré de l'écart-type. Nous emploierons fréquemment les expressions de "variance sociale" et de "variance naturelle" (plutôt que de variance des contraintes naturelles et distribuées. . .) par la suite, pour référer à l'"étalement" des distributions gaussiennes précédentes.

### Distribution liée au jeu de contextes

Cette troisième et dernière distribution va se construire de façon isomorphe à la précédente, en remplaçant l'ensemble des agents et de leurs liens par l'ensemble des contextes et de leurs relations. Il est possible de calculer pour chaque item **k** dans un contexte **C<sub>i</sub>** une **dévi**  
**ation contextuelle optimale** grâce à la même fonction  $d_s^l$  que précédemment, mais appliquée non plus à **RI** et aux fréquences d'un item chez différents agents, mais cette fois à **RC** et aux fréquences d'un item dans les différents contextes en jeu :

$$\mathbf{d}_c(f_k, C_i) = \sum_{j=1, j \neq i}^q d_s^l(k, C_j)$$

A partir de ce calcul, il est possible de calculer une troisième distribution centrée sur la déviation contextuelle optimale, avec une dernière variance dite "contextuelle" définie par l'utilisateur.

### Intégration des différentes distributions

Afin de déterminer la distribution finale des changements possibles pour un item, il nous reste à intégrer les différentes distributions précédentes.

Afin de combiner les distributions, nous allons simplement les sommer, et normaliser le résultat pour garder une intégrale des qualités de la distribution égale à 1. Ce calcul permet d'obtenir

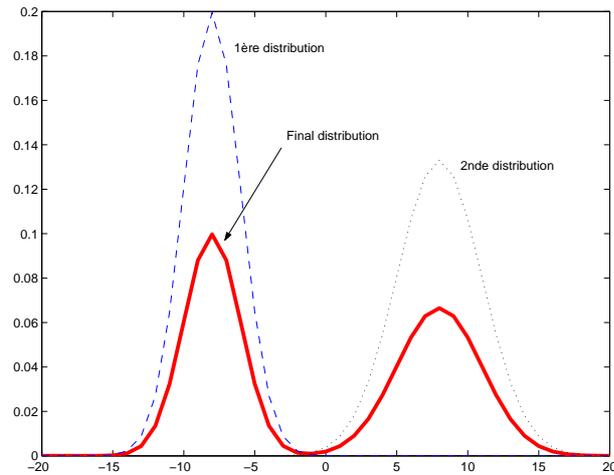


FIG. 6.3 – Somme normalisée de deux Gaussiennes présentant deux pics bien distincts

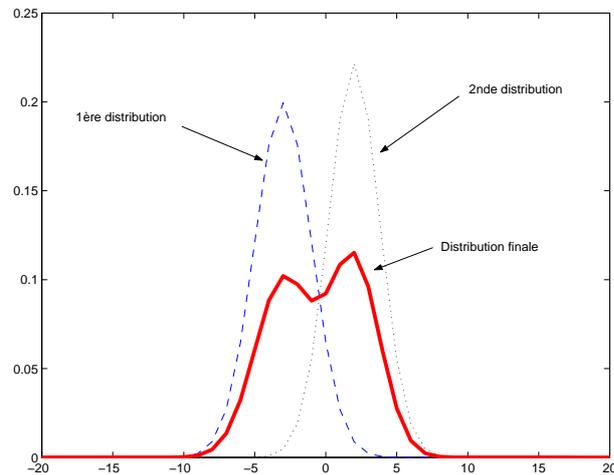


FIG. 6.4 – Somme normalisée de deux Gaussiennes présentant deux pics très rapprochés

une distribution finale qui prenne en compte les tendances des distributions correspondant aux différentes contraintes. Selon les caractéristiques de celles-ci, la distribution finale présentera un ou plusieurs “pics” plus ou moins rapprochés, comme le montrent les figures 6.3 et 6.4 pour deux sous-distributions.

Grâce à la combinaison des différentes gaussiennes, il est assez intuitif d’envisager comment les différentes contraintes internes ou externes au “système langage” pourront pousser l’évolution dans un sens ou dans l’autre. Ainsi, les contraintes sociales pourront à certains moments aller dans le sens des contraintes structurelles, et dans d’autres cas s’y opposer.

Différents paramètres peuvent jouer sur le système. Tout d’abord, jouer sur les variances des différentes distributions va permettre d’envisager pour chaque jeu de contraintes le test des différents scénarios détaillés plus haut. Il est en effet possible d’étendre leur principe, détaillé pour les contraintes naturelles, à l’ensemble des distributions.

Il est également possible d'adopter une somme pondérée pour combiner les différentes distributions, ce qui permet d'accorder plus ou moins d'importance aux différentes contraintes. Nous pourrions ainsi envisager par exemple des cas où l'évolution du système d'un agent est très liée aux contraintes sociales, et très peu aux contraintes internes.

### Evolution de la structure d'un système linguistique par mutations

A partir de la distribution globale établie précédemment, il nous reste à préciser comment les fréquences peuvent évoluer dans le système.

Comme nous l'avons déjà proposé plus haut, l'idée est donc de transformer la distribution de qualité des variations de fréquences potentielles pour chaque item pour la relier à la probabilité que telle ou telle variation se produise réellement à un moment donné. Les mutations devraient ainsi être définies de telle sorte que pour un nombre de tirages stochastiques assez important, la probabilité d'une variation corresponde à sa qualité dans la distribution. Le procédé est le suivant :

**Tirages de valeurs selon une distribution Gaussienne.** Soit  $f$  une fonction de distribution gaussienne (encore appelée fonction de densité normale) d'écart-type 1 et de moyenne 0 :

$$\forall x \in [-\infty, +\infty], \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

On désigne par  $\Psi$  la fonction de distribution cumulative normale. Cette fonction, qui correspond à la valeur cumulée de la fonction précédente selon  $x$ , se définit de la façon suivante :

$$\forall x \in [-\infty, +\infty], \quad \Psi(x) = \int_{-\infty}^x f(t) dt$$

La fonction de distribution cumulative permet de savoir pour une valeur  $x$  donnée le pourcentage de la distribution qui se trouve entre  $-\infty$  et  $x$ . Il n'existe pas de formule analytique de cette fonction, mais de nombreuses approximations de bonne qualité existent. La suivante est donnée par exemple par [Bagby, 1995] :

$$\forall p \in [0.0, 0.5], \quad \Psi'(x) = \frac{1}{2} \times \sqrt{1 - \frac{1}{30} \left( 7e^{-\frac{x^2}{2}} + 16e^{-x^2(2-\sqrt{2})} + (7 + \frac{1}{4}\pi x^2)e^{-x^2} \right)}$$

Désignons enfin par  $\Phi$  la fonction inverse de  $\Psi$ , c'est à dire la fonction inverse de la fonction de distribution cumulative normale. Cette fonction fait correspondre à l'intervalle  $[0, 1]$  l'intervalle  $[-\infty, +\infty]$ . Afin de produire des valeurs qui respectent en moyenne la distribution initiale  $f$ , il suffit dès lors d'effectuer un tirage aléatoire  $\pi$  entre 0 et 1, et de calculer la valeur de la fonction  $\Phi$  pour ce point.

Il n'existe pas de forme analytique de la fonction  $\Phi$ , mais comme pour  $\Psi$ , de très bonnes approximations existent [Abramowitz and Stegun, 1964].

Les résultats précédents concernent tous une fonction gaussienne d'écart-type 1 et de moyenne 0. Pour recourir à une fonction Gaussienne de moyenne  $\bar{X}$  et d'écart-type  $\sigma$ , Il suffit de transformer le résultat de la fonction  $\Phi$  en le multipliant par  $\sigma$ , puis en lui ajoutant  $\bar{X}$ .

**Tirages de valeurs selon une composition de distributions gaussiennes.** L'ensemble des formules précédentes concerne le cas d'une unique distribution normale, alors que notre distribution finale est elle la somme éventuellement pondérée de plusieurs gaussiennes. Alors qu'il était facile dans le cas précédent de recourir aux approximations de la fonction inverse de la distribution cumulative normale, nous ne pouvons ici que nous appuyer sur les approximations pour la fonction de distribution cumulative normale, et déterminer l'inverse de cette fonction selon un algorithme de calcul.

Soient  $(\overline{X}_1, \sigma_1)$ ,  $(\overline{X}_2, \sigma_2)$  et  $(\overline{X}_3, \sigma_3)$  les trois couples de moyenne et d'écart-type pour les trois gaussiennes  $\mathbf{f}_1$ ,  $\mathbf{f}_2$  et  $\mathbf{f}_3$ , et  $\alpha_1$ ,  $\alpha_2$  et  $\alpha_3$  leur poids relatif dans la somme finale. Nous définissons ensuite trois fonctions ( $i = 1, 2, 3$ ) représentant chacune la transformation pour normaliser une des trois Gaussiennes :  $\delta_i = \frac{x - \overline{X}_i}{\sigma_i}$  (ces fonctions ne rentrent pas dans le calcul de l'intégration ci-dessous, car elles appliquées de façon indépendante avant celui-ci).  $\mathbf{f}$  désigne la fonction gaussienne normale d'écart-type 1 et de moyenne 0.

Nous pouvons alors calculer de façon approximative la fonction de distribution cumulée pour la somme pondérée des trois gaussiennes de la façon suivante :

$$\forall x \in [-\infty, +\infty], \quad (6.1)$$

$$\Psi'_t(x) = \frac{1}{\alpha_1 + \alpha_2 + \alpha_3} \left( \int_{-\infty}^x (\alpha_1 \times f_1(t) + \alpha_2 \times f_2(t) + \alpha_3 \times f_3(t)) dt \right) \quad (6.2)$$

$$= \frac{1}{\alpha_1 + \alpha_2 + \alpha_3} \left( \int_{-\infty}^x (\alpha_1 \times f \delta_1(t) + \alpha_2 \times f \delta_2(t) + \alpha_3 \times f \delta_3(t)) dt \right) \quad (6.3)$$

$$= \frac{1}{\alpha_1 + \alpha_2 + \alpha_3} (\alpha_1 \times \Psi' \circ \delta_1(t) + \alpha_2 \times \Psi' \circ \delta_2(t) + \alpha_3 \times \Psi' \circ \delta_3(t)) \quad (6.4)$$

A partir de cette fonction monotone croissante, un algorithme assez simple permet de calculer la fonction inverse (donc de façon non analytique) pour toute valeur comprise entre 0 et 1. En travaillant sur un intervalle raisonnable (même si les gaussiennes peuvent être de moyennes différentes, l'ensemble des valeurs où la distribution générale est significative est restreint), une approximation à une précision suffisante (définie par un paramètre  $\varepsilon$ ) se fait aisément par **dichotomie**. Le nombre de pas de l'algorithme dépendra de cette précision requise. Notons ici que nous nous limiterons en fait dans nos programmes à des mutations fréquentielles d'au plus 0.05, ce qui conduira ici à un espace de recherche pour l'algorithme de dichotomie situé entre -0.05 et 0.05.

A chaque pas de temps, un calcul de la distribution de chaque item sera effectué, pour chaque individu et chaque contexte le cas échéant. Un premier tirage aléatoire  $\pi$  donnera un nombre entre 0 et 1, et par application du processus de dichotomie, la variation correspondante selon la courbe de distribution générale sera déterminée. Un seconde tirage  $\pi'$  donnera alors un nombre compris entre une valeur  $\mathbf{m}_r$ , comprise entre 0 et 1, et 1. Si ce nombre est inférieur à la valeur de la distribution générale pour la variation déterminée, alors cette variation deviendra effective, sinon il n'y aura pas de modification de la fréquence pour ce pas de temps.

Ce système à double tirage aléatoire permet de respecter en moyenne le fait qu'un changement de plus grande qualité se produit plus souvent par rapport aux autres changements à un instant donné, et plus fréquemment dans le temps.

A un temps  $t$ , l'ensemble des items sera ainsi testé, et une liste de variations à apporter sera ainsi déterminée. Après le test de tous les items, les fréquences de ceux-ci seront mises à jour selon la liste de variations, avant de passer au pas de temps suivant et réitérer l'opération. Ce mécanisme en deux temps permet d'éviter un effet d'ordre dans les changements des fréquences.

## Structures sociales et géographiques

Se reposer sur la notion de réseau social, avec des liens de forces différentes est un premier point, et il est bien sûr possible de les créer de multiples manières dans une population d'agents initialement non structurée. Afin de faciliter l'analyse et de pouvoir considérer plusieurs cas de figures assez généraux, il est souvent utile de recourir à des topologies particulières comme support des liens entre les individus.

Dans nos expériences, nous avons le plus souvent utilisé des réseaux complets (chaque agent est connecté à tous les autres) formés de liens égaux entre tous les agents.

Une seconde topologie qui se dégage assez naturellement et qui a déjà été utilisée dans de nombreuses études est celle d'une grille d'agents. Cette grille peut-être torique ou non : le passage d'une grille pourvue de bords à une grille dépourvue de limites s'effectue simplement par la prise en compte sur un bord des agents situés sur le bord opposé (on "recolle" les bords opposés).

Chaque intersection (ou chaque case) de la grille contient donc un agent. Il est possible (pour les modèles qui nous intéressent) d'utiliser la topologie de la grille de deux façons : pour définir une distance sociale comme l'a fait par exemple Nettle [Nettle, 1999c], ou pour définir une distance spatiale, comme il est courant de le faire dans de nombreuses études informatiques [Kirby, 1997] (voir chapitre 3). Dans les deux cas cités, seul un des deux aspects, soit l'aspect social, soit l'aspect spatial, est pris en compte, avec parfois l'idée implicite que l'un des aspects est subsumé par le second.

Il est possible de bien séparer et étudier les deux aspects précédents en envisageant deux topologies distinctes "superposées". La possibilité la plus simple est l'utilisation de la topologie en grille pour l'aspect spatial, et la superposition des liens sociaux sur cette grille (le traitement de la situation inverse est totalement symétrique, mais intuitivement plus difficile à se représenter). Bien que nous ne les ayons pas abordés et contrastés en détail par la suite, il est assez facile de spécifier différents scénarios examinant différentes configurations sociales et spatiales :

- une situation où les liens sociaux sont répartis de façon aléatoire sur la grille spatiale. La distribution des forces des liens peut varier ;
- une situation où les voisins spatiaux sont reliés par des liens forts, avec des liens très faibles sur de plus grandes distances (voire des liens nuls) ;
- en contraste, une situation où les liens à longue distance sont renforcés.

Nous présenterons toutefois par la suite une rapide expérience sur la dispersion spatiale de liens sociaux positifs.

### 6.1.4 Quelques remarques en guise de transition

Le modèle précédent possède selon nous deux caractéristiques intéressantes : la première est que les contextes linguistiques et ceux formés par les individus sont totalement interchangeables. Si la topologie des liens entre les contextes peut-être fort différente de celle des liens entre les

individus, le mécanisme de caractérisation des variations possibles pour un item à un instant donné est le même dans les deux cas.

Cette structure du modèle est pour nous le moyen d'illustrer de façon forte la bidirectionnalité des changements, lors par exemple du phénomène de diffusion lexicale. Si l'on se place du "point de vue" d'un changement, il existe deux directions de propagation : dans l'ensemble des contextes, et dans la communauté d'individus. Selon les conditions dans chacune de ces directions, différentes dynamiques d'évolution pourront être observées.

Le second point est que la généralité du modèle, en particulier à l'aide des différentes distributions de qualité des variations potentielles, permet d'envisager *a priori* différents types de changements mentionnés dans la littérature, à savoir les changements dits "internes" ("*internally motivated*") et les changements "sociolinguistiques" ("*externally motivated*"). Nous espérons pouvoir présenter un cadre unifié pour ces différentes dynamiques d'évolution, et en rendre compte grâce aux variations des paramètres du modèle, en particulier au niveau des réseaux sociaux.

Les premières remarques sur la bidirectionnalité des changements ne seront pas traitées dans ce présent travail. Notre objectif est plus ici d'étudier les croisements entre les contraintes sociales et les contraintes naturelles, et nous ne détournerons que peu de ce but pour étudier les phénomènes précédents.

Notre ligne de conduite sera la construction de plusieurs séries d'expériences pour étudier l'impact des différents paramètres du modèle. Nous complexifierons notre point de vue au fur et à mesure de nos expériences. En suivant la démarche de Nettle énoncée au chapitre 3, nous tenterons de déterminer les conditions dans lesquelles les trajectoires d'évolution des systèmes linguistiques abstraits correspondent à la réalité des phénomènes linguistiques.

## 6.2 Etude des dynamiques d'évolution d'un système linguistique

### 6.2.1 Première série d'expériences : impact de la distribution des variations fréquentielles sous contraintes naturelles et distribuées

#### Contraste des cas extrêmes et intermédiaire de l'impact de la variance naturelle sur les évolutions linguistiques

Un des aspects fondamentaux du modèle est le recours à une distribution de la qualité des variations fréquentielles possibles en fonction des contraintes. Une première étude à envisager est de déterminer les dynamiques d'évolution des trois scénarios que nous avons détaillés dans la première partie de ce chapitre : évolution par optimisation locale des contraintes, évolution stochastique sans prise en compte des contraintes, et évolution selon une distribution des changements. Les deux premiers cas correspondent en fait respectivement à un écart-type nul ou infini de la distribution.

Dans cette expérience, nous envisageons un seul système linguistique (par exemple une langue considérée comme homogène ou un idiolecte), composé d'un seul contexte. Dans ce contexte, deux items voient leur fréquence évoluer de façon continue entre 0 et 1, sous l'effet d'événements contingents et sous un jeu de contraintes naturelles et distribuées, selon une distribution de probabilités.

Pour chaque paire de fréquences des deux items, la fonction d'énergie va permettre de déterminer l'adéquation aux contraintes de cette configuration. Un seul contexte inclusif est mis en jeu. Afin de disposer d'un paysage présentant plusieurs puits d'énergie, et des barrières d'énergie

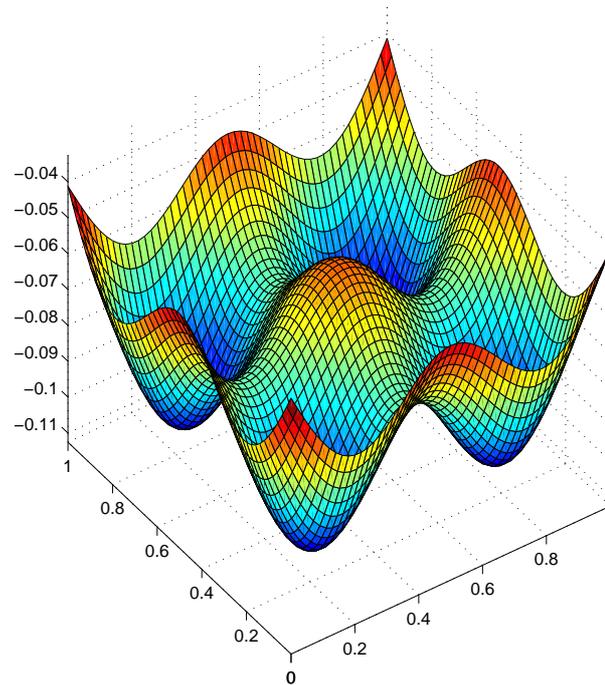


FIG. 6.5 – Paysage énergétique défini par la fonction  $f$

entre ces puits, la fonction considérée est la suivante :

$$\begin{aligned} \forall (x, y) \in [0, 1]^2, f(x, y) = & 0.1 \times \left( e^{-\frac{(x-0.2)^2}{0.3^2} - \frac{(y-0.2)^2}{0.3^2}} \right. \\ & + e^{-\frac{(x-0.2)^2}{0.3^2} - \frac{(y-0.8)^2}{0.3^2}} \\ & + e^{-\frac{(x-0.8)^2}{0.3^2} - \frac{(y-0.2)^2}{0.3^2}} \\ & \left. + e^{-\frac{(x-0.8)^2}{0.3^2} - \frac{(y-0.8)^2}{0.3^2}} \right) \end{aligned}$$

Le paysage énergétique de cette fonction est représenté sur la figure 6.5. Chaque exponentiel de la fonction correspond en fait à l'un des puits observés sur la figure.

Une implémentation du modèle d'évolution linguistique a été réalisée en C++ grâce à la plateforme LEMMING. Dans cette expérience, seule une partie des composantes programmées est réellement à l'oeuvre, puisqu'une seule distribution de qualité entre en jeu. L'histoire évolutive d'un système lors d'une simulation est préservée dans un fichier texte, sous la forme de la succession des paires de fréquences au cours du temps. Le tracé de ces paires sous Matlab sur le paysage énergétique permet de suivre graphiquement l'évolution du système, et de l'envisager en lien avec la topographie du système.

A partir d'une configuration initiale dans l'espace des paramètres, le système évolue pendant un certain nombre d'unités de temps (plusieurs dizaines de milliers le plus souvent).

Le premier cas considéré est celui d'une distribution d'écart-type nul (en fait extrêmement faible dans nos programmes, ce qui explique que la courbe ne soit pas rectiligne sur la figure,

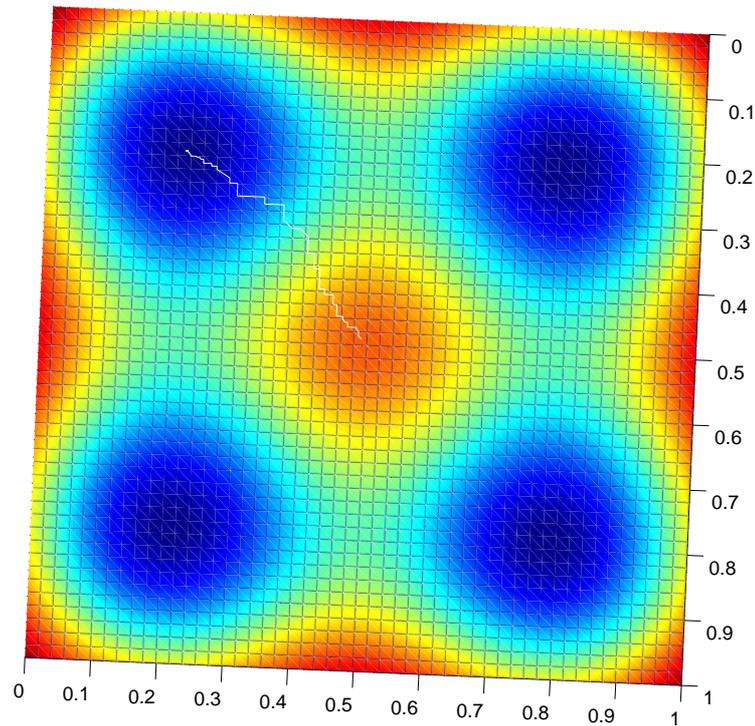


FIG. 6.6 – Evolution du système sous contrainte par optimisations locales

comme elle devrait l'être dans le cas théorique). Toute transformation qui se produit dans le système va dans le sens d'une meilleure adéquation aux contraintes. La fréquence même d'apparition de ces transformations est proportionnelle à leur qualité.

On constate que le système suit la ligne de plus grande pente du paysage à partir de son point de départ, et rejoint au bout d'un certain temps le minimum local dont le bassin d'attraction englobait sa position initiale. Après avoir atteint cet optimum, le système se stabilise. Dans les calculs, les dérivées partielles de la fonction énergétique sont nulles, ce qui ne permet plus de modifier les valeurs des deux fréquences. En effet, les distributions des déviations fréquentielles pour les deux items sont centrées sur 0, et la nullité de leur écart-type empêche dès lors tout changement. La figure 6.6 reproduit un exemple d'évolution du système.

Le second scénario correspond à une non prise en compte des contraintes ; l'écart-type de la distribution peut être considéré comme infini (dans le programme, il prend en fait une valeur positive très importante). On observe une évolution stochastique du système au cours du temps, sans respect du relief du paysage énergétique. Le système ne se stabilise pas, et bien que les mutations soient aléatoires au cours du temps, la probabilité qu'un système reste fixe pendant une période de temps diminue rapidement avec l'accroissement de la durée de cette période. En outre, si le système est instable, il n'est cependant pas possible de parler de fluctuations du système autour d'un état particulier de l'espace des paramètres. La figure 6.7 fournit un exemple d'évolution stochastique.

Le dernier scénario repose sur une distribution d'écart-type moyen, ni trop important ni trop faible, afin d'obtenir en moyenne des transformations allant dans le sens d'une meilleure adéquation aux contraintes, mais sans exclure la possibilité d'autres évolutions allant en direction

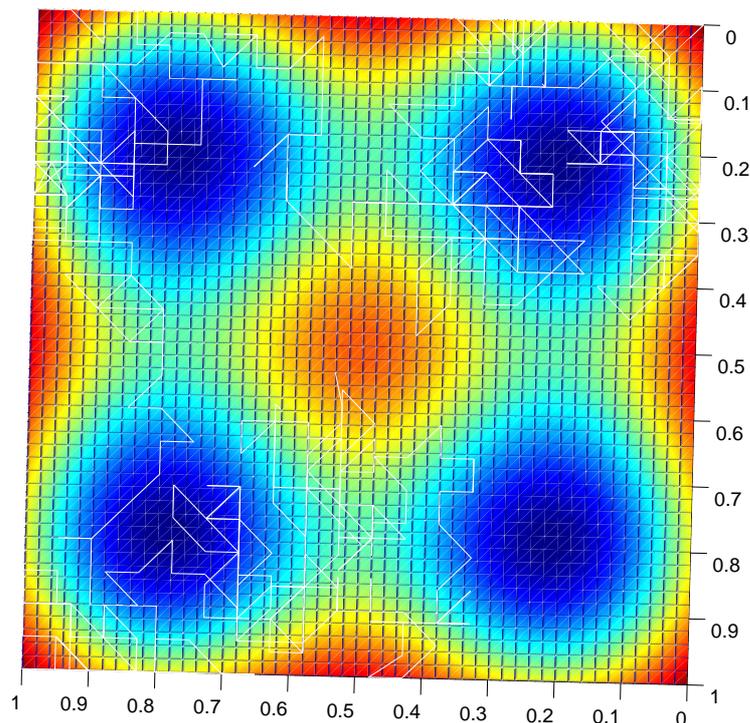


FIG. 6.7 – Evolution du système en l’absence de contraintes

opposée. Les figures 6.8 et 6.9 permettent de mieux se représenter la situation : le système fluctue pendant un certain temps autour d’un attracteur (un minimum local), avant de franchir une barrière énergétique et d’aller fluctuer auprès d’un second. Les raisons de cette évolution sont les suivantes : les évolutions stochastiques biaisées ont tendance à maintenir le système près de l’optimum local, mais une partie plus faible des mutations a tendance au contraire à l’éloigner de ce minimum ; plus les pentes du bassin d’attraction sont faibles, plus nombreuses seront les mutations qui tendront à éloigner le système du minimum. Par le biais du mécanisme stochastique, il est possible à une certaine période que plusieurs mutations divergentes se produisent successivement et emmènent le système jusqu’à la frontière du bassin d’attraction de l’optimum. Si une dernière mutation parvient à faire quitter le bassin au système, celui-ci peut alors être capté par un nouvel attracteur et se diriger vers lui par le biais de mutations en accord avec la pente de son bassin d’attraction. Une nouvelle période de fluctuations commence alors autour de cet attracteur, jusqu’à ce que la situation précédente se reproduise.

Notons ici qu’il n’est pas possible de prévoir quand le système basculera d’un attracteur vers un autre. Si l’on se remémore les idées de Ehala énoncées au chapitre 2, les mutations stochastiques représentent des événements contingents, et rendent dès lors imprévisible l’évolution du système. Toutefois, comme cela est souvent le cas avec des mécanismes aléatoires, un système ne peut conserver le même comportement indéfiniment : la probabilité qu’une pièce tombe toujours du côté face décroît non linéairement avec le nombre de tentatives, et la possibilité qu’une pièce tombe indéfiniment du même côté est nulle. De façon analogue ici, la probabilité que le système reste indéfiniment proche du même attracteur est nulle si les barrières énergétiques qui l’entourent peuvent être franchies. Toutefois, selon la taille du bassin d’attraction et l’inclinaison

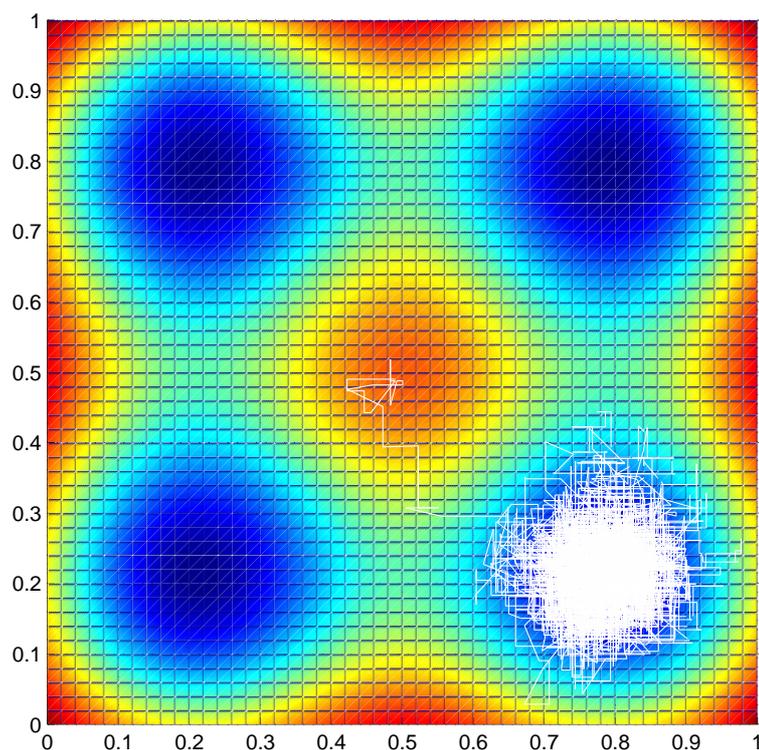


FIG. 6.8 – Evolution selon une distribution des changements sous contraintes après 10,000 pas de temps

de ses pentes, le temps moyen pour échapper à cet attracteur sera plus ou moins long. Seul le comportement asymptotique du système peut-être déterminé selon une certaine probabilité.

### Définition d'indicateurs supplémentaires pour l'évolution du système

Dans les cas précédents et les suivants, l'espace des paramètres est de dimension 2, ce qui permet de visualiser aisément les évolutions d'un système. Toutefois, ceci n'est plus possible pour un nombre de dimensions plus important. Il est cependant possible de définir un certain nombre d'indicateurs qui permettent de détecter les périodes de stabilité et d'instabilité d'un système quel que soit le nombre de dimensions. Les scénarios en dimension 2 ou 3 sont alors utiles, puisqu'ils permettent de tester la pertinence des indicateurs. Le problème est que, comme dans nombre de problèmes d'optimisation en mathématiques, la topographie du paysage énergétique est inconnue est qu'il n'est pas possible de s'appuyer sur la position des attracteurs pour calculer si un système est à proximité d'eux ou non.

Une méthode possible est d'essayer de supprimer en partie les fluctuations du système autour d'un minimum du paysage énergétique. Pour cela, il est possible de calculer le barycentre des positions du système pendant un certain laps de temps  $\mathbf{T}$ . Si le système demeure dans le bassin d'un attracteur au cours de cette période et si celle-ci est suffisamment longue, alors le barycentre calculé doit être proche de l'attracteur. Il faut toutefois nuancer ce point par le fait que cette proximité est sujette au relief plus ou moins plat du bassin de l'attracteur (plus le bassin est plat, plus la période de temps pour le calcul du barycentre devra être importante

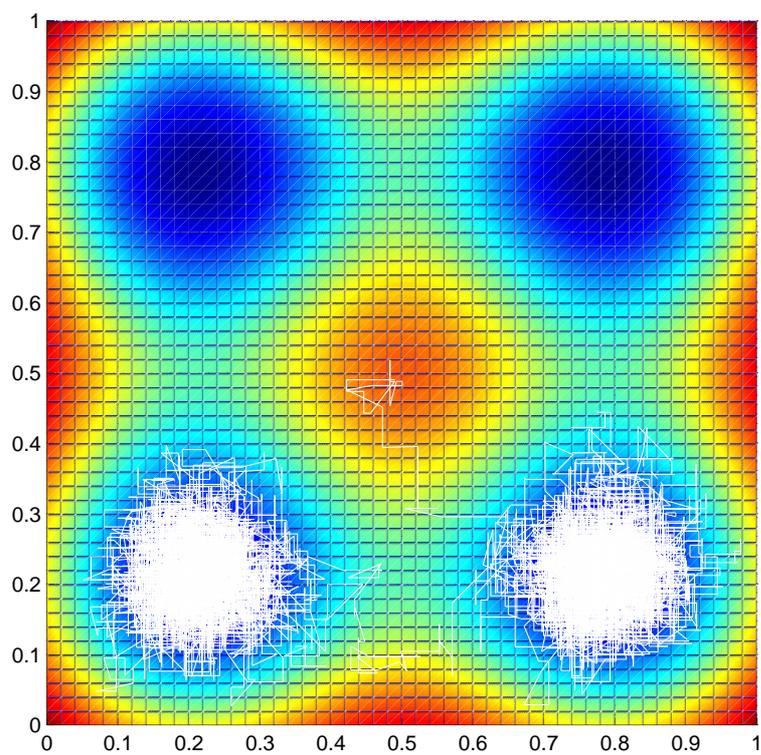


FIG. 6.9 – Evolution selon une distribution des changements sous contraintes après 20,000 pas de temps

pour que celui-ci soit proche de l'attracteur).

Nous pouvons définir un indicateur de *stabilité structurelle* de la façon suivante. Soit  $\mathbf{T}$  une période de temps que nous divisons en deux parties égales. Soit  $\mathbf{B}_1$  le barycentre des positions du système pour la première période, et  $\mathbf{B}_2$  celui de la seconde période. Nous calculons alors la distance entre les deux barycentres.

Considérons le cas où le système franchit une barrière énergétique et passe d'un premier bassin d'attraction à un second à la moitié de la période  $\mathbf{T}$ . Dès lors, la distance entre le premier et le second barycentre sera plus importante que dans le cas où le système reste dans le même bassin d'attraction au cours de la période  $\mathbf{T}$ . Si l'on représente la distance entre les barycentres au cours du temps, une hausse significative de celle-ci traduira un changement de bassin d'attraction.

Encore une fois, l'accroissement de la distance est fonction de la topographie des bassins d'attraction. Il sera d'autant plus faible que ces bassins seront peu marqués et séparés par des barrières énergétiques de faible épaisseur.

Un des moyens de corriger ce problème est en fait d'analyser plus finement l'évolution de la distance entre les barycentres successifs. Cette approche permet en fait de mettre au jour des informations intéressantes à propos d'un paysage énergétique inconnu.

Il convient tout d'abord de déterminer quelle période de calcul  $\mathbf{T}$  permet de dégager le plus d'informations à partir de la suite des positions du système linguistique dans l'espace des possibles. Pour une topographie donnée, une valeur trop petite est sujette aux fluctuations rapides du système, tandis qu'une valeur trop importante risque d'entraîner la perte d'informations si les variations du système sont rapides.

Ensuite, il convient de détecter les changements significatifs dans l'évolution de la distance, afin de pouvoir identifier les transitions entre les différents bassins d'attraction. Une fois cette opération effectuée, l'examen de la variation de la distance à l'intérieur d'un bassin d'attraction permet de dégager quelques informations sur la topographie de celui-ci, à condition que la période pendant laquelle le système reste dans ce bassin soit assez importante. En particulier, il est possible d'estimer au mieux la position de l'attracteur par le calcul du barycentre sur la période entière de stabilité.

Nous voyons également apparaître la possibilité d'une application récursive des principes d'analyse à l'intérieur d'une zone de stabilité. En choisissant les meilleures paramètres d'étude pour cette dernière, il sera possible d'avoir une idée plus précise de sa topographie, détecter des minima locaux. . .

Une façon de définir un bon indicateur général de stabilité structurelle, dont les valeurs dépendent moins des distances entre les attracteurs, est de diviser la distance entre les barycentres par la moyenne des deux écarts-types des distributions établis pour les deux périodes successives de la période  $\mathbf{T}$ .

Les courbes 6.10, 6.11 et 6.12 représentent successivement les variations d'énergie d'un système et les deux indicateurs de changement (relatif et absolu) définis plus haut. Le paysage est le même que celui défini précédemment.

### **Impact de la variance de la distribution des transformations fréquentielles sur le nombre de changements de bassin d'attraction**

Comme souligné plus haut, les trois scénarios déjà étudiés correspondent à différentes valeurs de la variance de la distribution des transformations possibles. Une question intéressante est d'examiner comment le nombre de changements, c'est à dire le nombre de passages d'un bassin

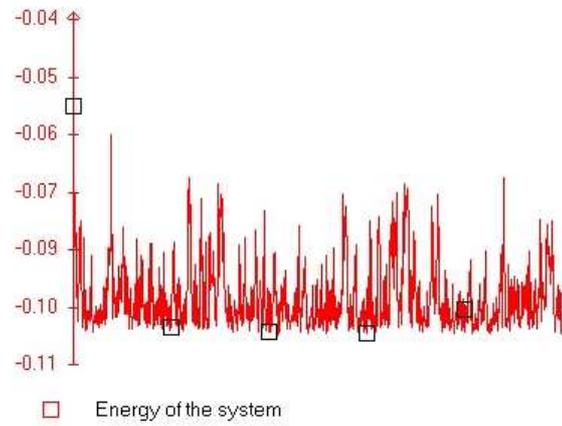


FIG. 6.10 – Evolution de l'énergie du système linguistique

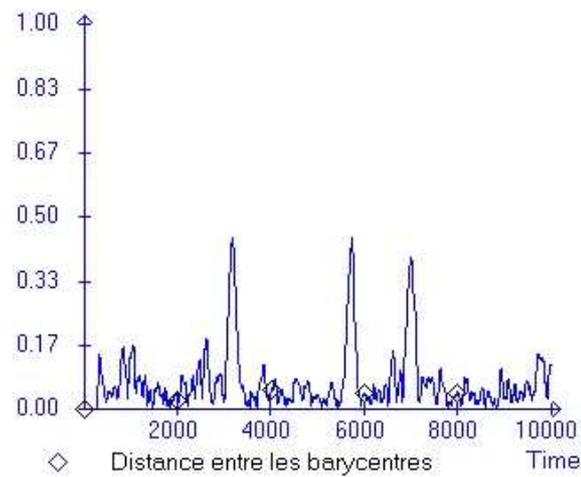


FIG. 6.11 – Un indicateur des changements d'un système en évolution : distance entre les barycentres du système pour deux périodes consécutives

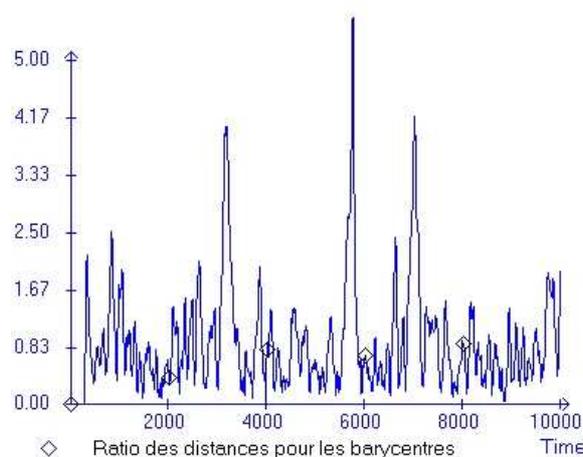


FIG. 6.12 – Evolution du ratio de la distance entre barycentres sur la moyenne des écarts-types des distributions des deux périodes temporelles consécutives considérées

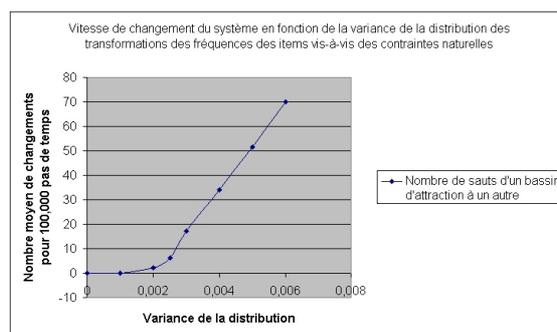


FIG. 6.13 – Evolution du nombre de changements en fonction de la variance de la distribution des transformations possibles

d'attraction à un autre, varie avec la valeur de la variance. En particulier, y a-t-il une valeur limite en-dessous de laquelle il n'est pas possible de quitter un bassin d'attraction ?

La courbe de la figure 6.13 répond aux questions précédentes, et met en évidence pour le système étudié une valeur seuil pour la variance proche de 0.001 sous laquelle le système ne peut quitter son bassin d'attraction initial. Elle montre également qu'au-delà de ce seuil, le nombre de changements par unité de temps est relié linéairement à la variance. Ceci ne peut toutefois être vrai que pour des valeurs de la variance assez faible, puisqu'une limite supérieure sera établie par la situation où le système se déplace aléatoirement sans respect du relief dans le paysage énergétique.

La valeur limite dépend de la profondeur des bassins d'attraction, et des barrières énergétiques à franchir pour passer de l'un à l'autre. Toutefois, le point intéressant ici est qu'une transition de phase en fonction de la variance existe d'une façon générale quelque soit la topographie du paysage énergétique. Ce résultat n'est pas intuitif, puisqu'on pourrait penser que toute barrière énergétique est franchissable si un nombre suffisant de mutations se succèdent dans une direction opposée à celle de l'attracteur.

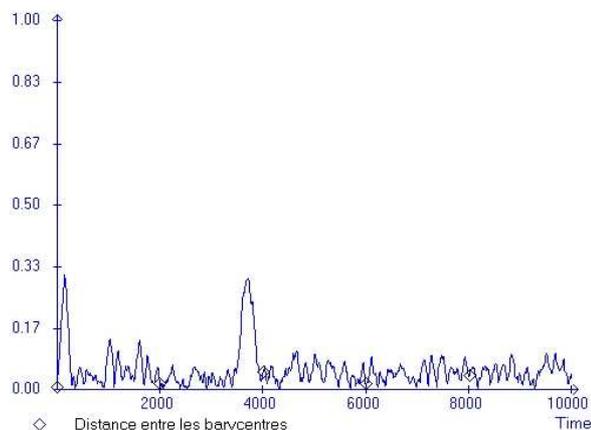


FIG. 6.14 – Evolution de la distance entre barycentres pour une variance égale à 0.002

Si l'on compare l'évolution de la distance entre les barycentres de deux périodes successives, on remarque qu'une variance plus importante entraîne non seulement plus de changements, mais également une distance plus importante entre les barycentres lors d'une période de stabilité structurelle (comparer les figures 6.14 et 6.15). Ceci s'explique logiquement par le fait que la variabilité est plus importante au sein du bassin d'attraction, et que ce phénomène se reflète au niveau des barycentres.

Si l'on analyse le résultat précédent en modifiant conceptuellement non pas la variance de la distribution, mais la topographie d'un bassin d'attraction, nous pouvons dire que plus ce dernier est pentu en direction de l'attracteur, plus un système reste proche de ce dernier, et plus les distances entre barycentres de périodes successives sont faibles (puisque les barycentres sont très proches de l'attracteur). Dès lors, étudier les valeurs de la distance entre les barycentres pour un bassin d'attraction donné permet de savoir si celui-ci est plutôt pentu et/ou de faibles dimensions, sans qu'il soit possible toutefois de trancher entre ces deux caractéristiques.

Avant de passer à l'étude de populations d'individus munis chacun de leur système linguistique (leur idiolecte), nous souhaitons présenter deux courtes expériences qui vont nous permettre de concrétiser certaines propositions énoncées au chapitre 4.

### Accroissement de la diversité par exploration de l'espace des possibles

Nous souhaitons mettre ici en valeur par un exemple simple l'accroissement au cours du temps de la diversité d'une famille de systèmes linguistiques à partir d'une situation initiale où les systèmes sont tous dans le même état structural. Nous allons supposer ici un ensemble de 10 systèmes linguistiques indépendants, et montrer comment une exploration stochastique du paysage va conduire à une réflexion progressive de l'espace des états possibles dans l'espace des états réels (voir chapitre 4).

Afin d'avoir un espace des possibles plus riche que celui défini plus haut, nous étendons l'espace des possibles afin de disposer d'un plus grand nombre de bassins d'attraction. La figure 6.16 représente cette transformation qui permet d'obtenir 16 bassins d'attraction.

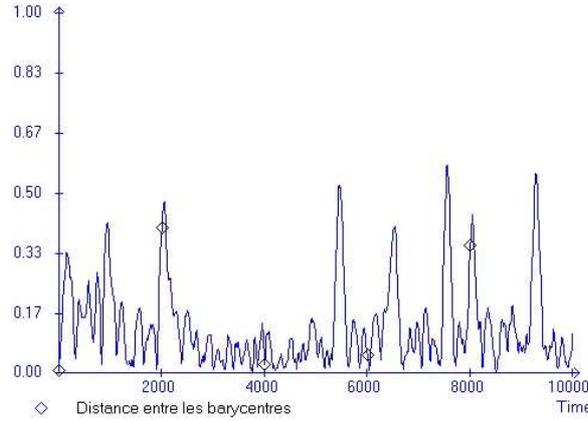


FIG. 6.15 – Evolution de la distance entre barycentres pour une variance égale à 0.006

Une simplification par rapport au scénario de polygenèse est ici que nous envisageons un état initial où l'ensemble des systèmes est déjà présent, et non pas une augmentation progressive du nombre de systèmes, ce qui a probablement été le cas dans la réalité avec la scission de groupes humains devenus trop importants.

Afin de pouvoir mesurer la diversité de l'ensemble des systèmes, nous reprenons un indicateur d'entropie mentionné dans [Oudeyer, 2002]. Si  $(x_1(t), x_2(t), \dots, x_n(t))$  sont les positions des différents systèmes à la date  $t$ , nous définissons tout d'abord pour un point quelconque  $x$  la fonction de densité de probabilité suivante :

$$p_n(x, t) = \frac{1}{n} \sum_{i=1}^n \frac{1}{2\pi\sigma} e^{-\frac{\|x-x_i(t)\|}{\sigma^2}}$$

La valeur de  $\sigma$  doit être définie de façon correcte par l'utilisateur pour bien capturer les caractéristiques de la distribution des  $(x_1(t), x_2(t), \dots, x_n(t))$ , mais un large intervalle de valeurs donne des résultats corrects. L'indicateur d'entropie est alors défini de la façon suivante :

$$entropy(t) = - \sum_{i=1} l p_n(X_i, t) \ln(p_n(X_i, t))$$

où les  $(X_1(t), X_2(t), \dots, X_l(t))$  sont les intersections d'une grille régulière qui recouvre le paysage énergétique.

La courbe 6.17 permet de suivre l'évolution de l'entropie et donc de la diversité du système. On remarque un accroissement initial de l'entropie, qui correspond à une répartition progressive des systèmes dans les différents bassins d'attraction. L'asymptote atteinte au bout d'environ 10,000 pas de temps correspond intuitivement à une distribution où chaque système se trouve dans un bassin différent. Toutefois, cette valeur limite est rarement atteinte puisque suivant les déplacements stochastiques des agents, il est probable qu'une majeure partie du temps, au moins deux agents partagent le même bassin d'attraction. La courbe 6.18 illustre quant à elle la

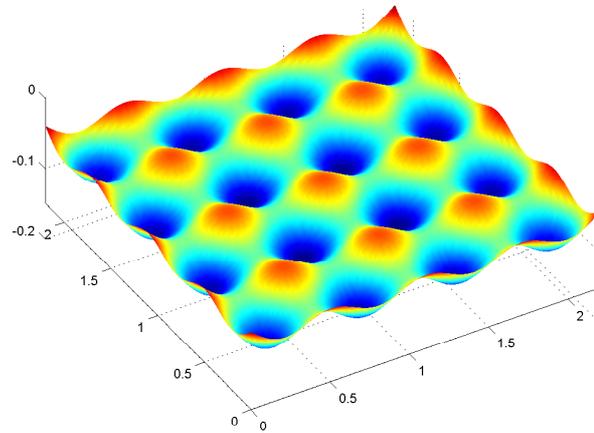


FIG. 6.16 – Un paysage énergétique avec seize bassins d'attraction

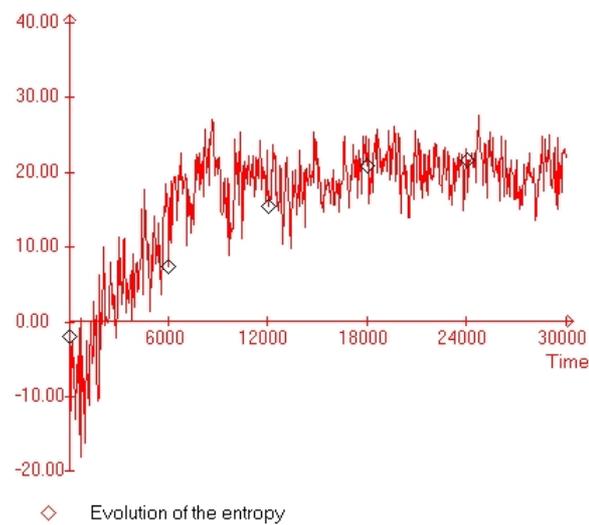


FIG. 6.17 – Evolution de l'entropie d'un ensemble de systèmes linguistiques lors de l'exploration de l'espace des possibles

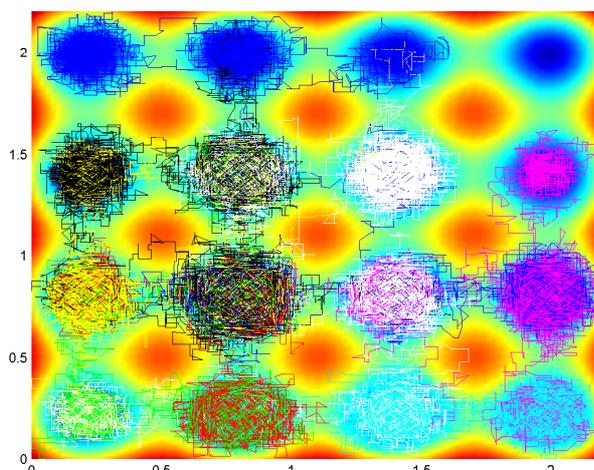


FIG. 6.18 – Vue de la diversité linguistique d'un ensemble de systèmes linguistiques après exploration de l'espace des possibles

situation atteinte après 30,000 pas de temps. On constate que l'exploration par dérive de l'espace des possibles a conduit à la découverte de 15 des 16 bassins du paysage.

Le problème de la mesure d'entropie utilisée est qu'elle n'est applicable qu'à des simulations mettant en jeu un nombre suffisant de systèmes. Si l'on ne considère que deux ou trois éléments, les variations entropiques sont trop faibles entre des configurations assez différentes, ce qui rend cette mesure peu pertinente.

### 6.2.2 Deuxième série d'expériences : Réseaux sociaux et évolution d'une communauté de locuteurs dans un paysage énergétique plat

Dans la première expérience, nous avons supposé un système linguistique isolé, ce qui pose le problème des interactions entre plusieurs systèmes, aussi bien celles de locuteurs dans une communauté isolée, que les contacts entre langues.

Le but de cette seconde série d'expériences sera d'étudier comment les liens sociaux viennent peser sur les évolutions linguistiques. Afin de procéder du plus simple au plus complexe, les simulations présentées dans cette série considèrent une absence de contraintes internes, ce qui se traduit par un paysage énergétique "plat". Seules les contraintes sociales sont ainsi à l'œuvre pour déterminer les évolutions linguistiques des systèmes.

#### Distance moyenne au barycentre et distance parcourue en moyenne par pas de temps

Afin d'étudier un macro-système composé de plus d'un système linguistique (nous dénommerons ces systèmes par le terme d'agent), nous allons recourir à plusieurs indicateurs que nous allons maintenant décrire.

Un premier indicateur repose sur la distance moyenne parcourue par pas de temps par les agents. Ceci permet d'apprécier la façon dont les liens sociaux viennent contraindre les évolutions des membres d'un réseau. Cette distance moyenne peut être assimilée au *rythme* ou à la *vitesse d'évolution* d'un système.

Il est également possible de se baser sur l'évolution du barycentre du macro-système. Pour cela, à chaque pas de temps, le barycentre des positions des sous-systèmes dans l'espace des paramètres est calculé. La succession de ces barycentres "synchroniques" au cours du temps permet d'observer la trajectoire "diachronique" du barycentre du système dans l'espace des paramètres.

Deux indicateurs peuvent être défini à partir du barycentre du macro-système. Un premier repose sur la distance moyenne d'un agent au barycentre. Ceci permet d'estimer approximativement comment les liens sociaux contraignent les agents à être proches ou au contraire éloignés les uns des autres. Cet indicateur peut donc être envisagé pour l'instant comme une mesure de la diversité du macro-système, même si l'absence de contraintes internes ne permet pas vraiment de parler de changements et de diversité au sens où nous les avons définis au chapitre 2.

Un second indicateur rend compte de la distance moyenne parcourue par le barycentre par unité de temps. Ceci permet cette fois de mesurer la façon dont la structuration du réseau social influence l'évolution du macro-système dans sa totalité, et renseigne sur le rythme d'évolution de ce dernier.

Il est possible de contraster la vitesse d'évolution d'un macro-système avec la vitesse moyenne des systèmes qui le composent (par exemple la rapidité d'évolution d'une langue par rapport à celle de ces idiolectes). Il en est particulier possible alors d'observer comme s'opère une certaine stabilisation du macro-système malgré les évolutions rapides de ces membres.

Enfin, le contraste entre la distance moyenne parcourue par le barycentre et celle parcourue par les agents permet de se faire une idée sur la façon dont s'opère une certaine stabilisation du macro-système malgré les évolutions de ces membres en raison de la structuration sociale.

### **Influence des liens sociaux sur l'évolution de deux locuteurs**

Une première étape pour évaluer l'importance des réseaux sociaux est d'évaluer l'influence de la valeur des liens sociaux entre deux locuteurs isolés, avant d'aborder des communautés plus importantes comme nous le ferons plus bas.

Les liens étudiés ici sont tous symétriques : le lien qui traduit l'influence du locuteur **A** sur le locuteur **B** est de même valeur que le lien qui représente l'influence du locuteur **B** sur le locuteur **A**.

Selon la valeur des liens entre deux locuteurs, leur trajectoire évolutive vont être plus ou moins voisines, ou au contraire plus ou moins éloignées dans l'espace des possibles.

Les figures 6.19, 6.20, 6.21, 6.22 et 6.23 permettent d'illustrer successivement les cas où les liens réciproques entre les deux agents sont : nuls, faiblement positifs, fortement positifs, faiblement négatifs et fortement négatifs. Il est assez aisé de constater que pour des liens positifs, les deux trajectoires sont d'autant plus proches que la valeur de ces liens est importante, et que pour des liens négatifs, les deux trajectoires seront au contraire plus distantes en cas de valeur (absolue) plus élevée. Dans le cas de liens nuls, les locuteurs se déplacent selon deux marches aléatoires indépendantes. Rappelons ici que si l'on se reporte à la définition de la déviation sociale optimale, la relation entre la force des liens et le comportement des locuteurs n'est pas linéaire, de même que celle entre la distance des locuteurs dans l'espace des possibles et ce même comportement. Ceci se traduit par exemple par le fait que plus la force des liens augmente, et moins l'augmentation de leur impact est importante. De même, des agents unis par des liens répulsifs se repoussent d'autant moins qu'ils sont distants dans l'espace des possibles.

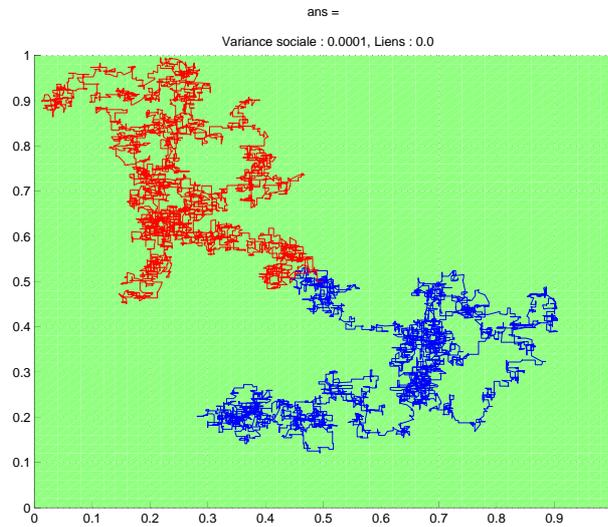


FIG. 6.19 – Evolution de deux locuteurs en l'absence de liens sociaux (variance sociale = 0.0001)

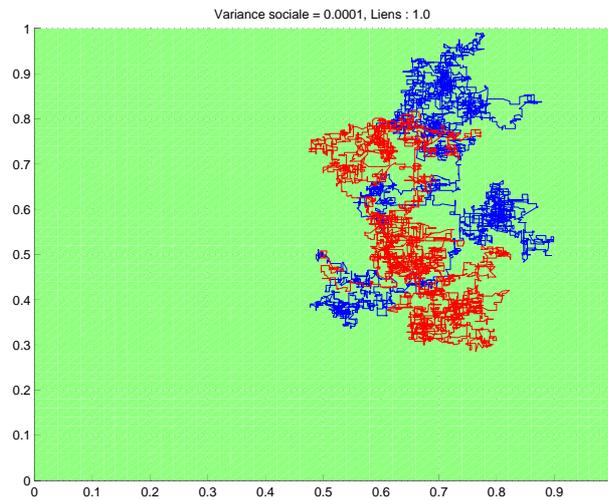


FIG. 6.20 – Evolution de deux locuteurs unis par deux liens positifs symétriques de valeur 1.0 (variance sociale = 0.0001)

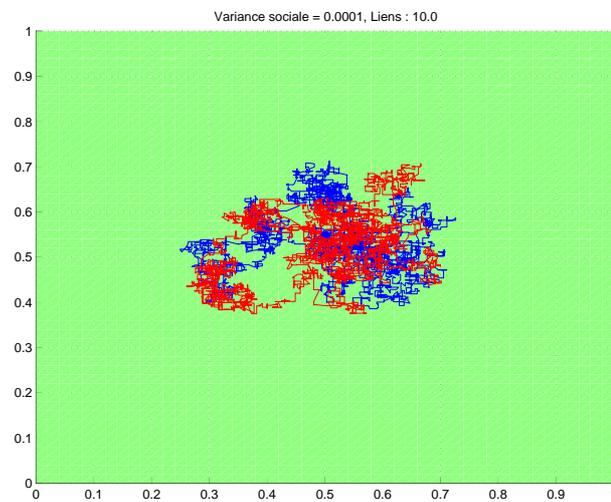


FIG. 6.21 – Evolution de deux locuteurs unis par deux liens positifs symétriques de valeur 10.0 (variance sociale = 0.0001)

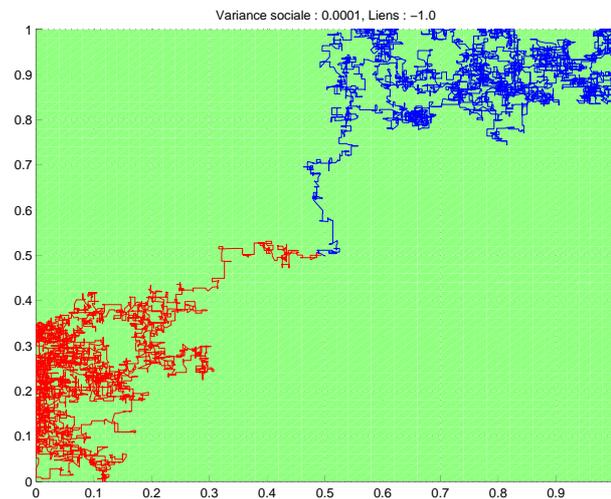


FIG. 6.22 – Evolution de deux locuteurs unis par deux liens négatifs symétriques de valeur -1.0 (variance sociale = 0.0001)

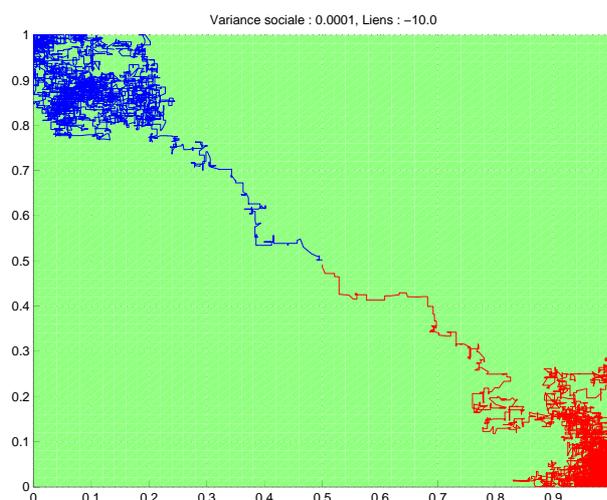


FIG. 6.23 – Evolution de deux locuteurs unis par deux liens négatifs symétriques de valeur -10.0 (variance sociale = 0.0001)

Il est possible de quantifier plus précisément la distance entre deux trajectoires évolutives en mettant à profit les indicateurs définis ci-dessus. La figure 6.24 représente ainsi la valeur de la distance moyenne au barycentre des deux agents en fonction de la valeur des liens qui les unissent, et ceci pour deux valeurs de la variance sociale (0.001 et 0.00001). On constate que cette distance se distribue selon la force des liens comme une fonction sigmoïde : ses variations sont pratiquement nulles pour une valeur absolue des liens importante, et une transition rapide se produit autour d'une valeur des liens nulle. La distance est importante pour des liens négatifs, et beaucoup plus faible pour des liens positifs, ce qui est naturel puisque des agents unis par des liens attractifs auront tendance à être proches, alors qu'une répulsion sociale se traduira par une plus grande distance entre les agents, et donc une plus grande distance moyenne au barycentre.

Notons ici que la valeur limite atteinte avec des liens négatifs et pour la variance sociale la plus faible correspond à une opposition presque maximale entre les deux agents, situés chacun près d'un angle de l'espace des paramètres de forme carrée (les agents occupent deux angles opposés). Le barycentre se trouve alors au centre de l'espace et la distance entre cette position et un angle du carré est égale à  $\sqrt{(0.5^2 + 0.5^2)} = 0.70$ . Pour une variance plus importante (0.001) et des liens toujours négatifs, la répulsion n'est plus aussi forte et les agents se déplacent à plus grande distance des coins opposés ; la distance moyenne au barycentre est augmentée en conséquence.

Plus généralement, pour la variance sociale la plus importante, les agents ne sont pas aussi contraints par les contraintes sociales que dans le cas de plus petite variance, et la distance moyenne au barycentre est plus faible pour les liens négatifs, et plus importante pour les liens positifs.

La figure 6.25 illustre quant à elle la distance moyenne parcourue par pas de temps par le barycentre et par les agents en fonction des liens sociaux. On constate que la distance parcourue par le barycentre est moins importante que celles des agents, ce qui traduit une certaine stabilité du macro-système vis à vis des évolutions de ces membres. Mais l'on observe surtout que les distances parcourues sont pratiquement constantes quelle que soit la valeur des liens entre les

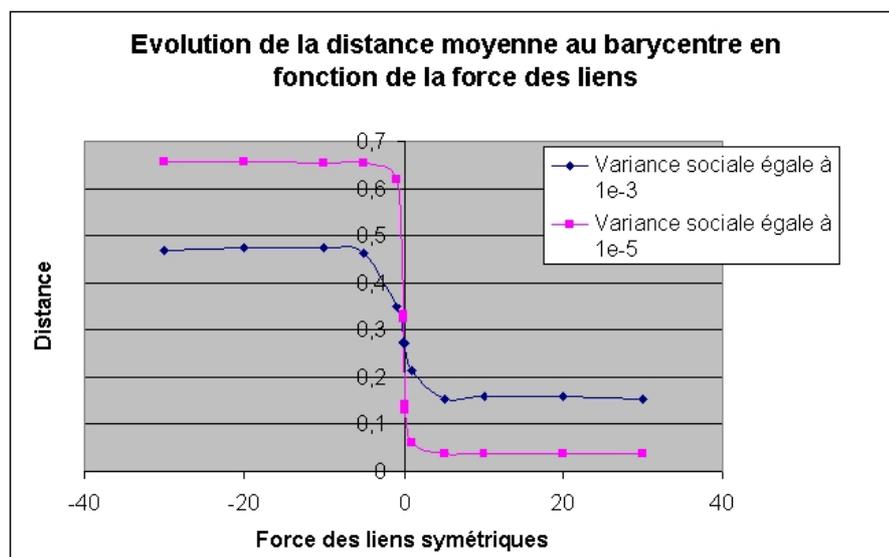


FIG. 6.24 – Evolution de la distance moyenne au barycentre en fonction des liens (symétriques) entre agents

agents. La distance moyenne parcourue par le barycentre et les agents n'est donc pas influencée *dans le cas de deux agents* par la force des liens qui les unissent (voir les cas à plus de deux agents plus bas). En fait, une observation très fine du graphique permet de déceler une très faible augmentation lors du passage de liens négatifs à des liens positifs, mais celle-ci est très faible par rapport aux valeurs des distances moyennes parcourues (de l'ordre de 1 ou 2 %).

On constate en outre que les distances moyennes parcourues par les agents et le barycentre sont plus faibles pour une variance plus petite. Ceci s'explique par le fait qu'une variance plus faible oblige les agents à être beaucoup plus congruents pour leur déplacements, et que leur capacité d'exploration de l'espace des possibles sera réduite en conséquence. Une réduction du rythme d'évolution des agents entraîne alors naturellement une diminution du rythme d'évolution du macro-système. Nous remarquons enfin que pour la plus faible variance, la vitesse d'évolution du barycentre est pratiquement égale à celle des agents. Ceux-ci sont si en accord sur leur déplacement que leur vitesse d'évolution s'assimile à celle du barycentre.

### Influence de la variance

Dans la première expérience précédente, nous avons contrasté les valeurs des indicateurs selon la force des liens sociaux et également deux valeurs de la variance sociale.

Il nous paraît intéressant d'examiner plus en détail l'impact de cette variance sociale sur les évolutions du système.

Comme nous l'avons fait plus haut, il nous paraît "pédagogique" de commencer cette courte étude en illustrant la problématique par deux figures présentant les trajectoires évolutives des deux agents. Ceux-ci sont liés dans l'expérience par des liens attractifs symétriques de valeur 10, et les deux variances sont égales à 0.001 et 0.0001. Les deux figures 6.27 et 6.26 permettent de constater que dans le cas d'une plus faible variance sociale, les trajectoires des deux agents sont beaucoup plus voisines et localisées spatialement que dans le cas d'une variance faible après le

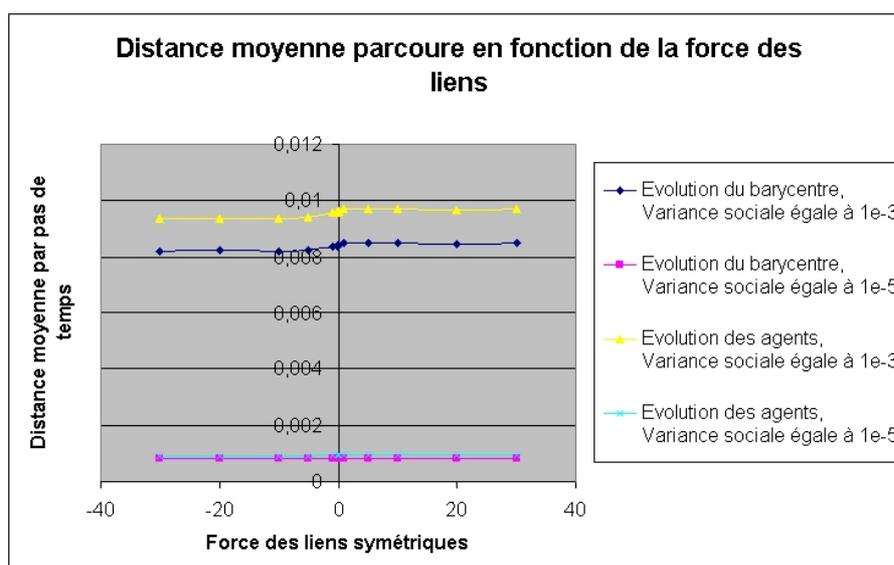


FIG. 6.25 – Evolution de la distance moyenne parcourue par les agents et le barycentre en fonction des liens (symétriques) entre agents

même nombre de pas d'évolution.

Afin de tester les impressions précédentes nées de l'observation des trajectoires évolutives des agents, nous avons étudié l'évolution des différents indicateurs en fonction de la variance sociale. Les valeurs de celle-ci varient entre  $1e-7$  et  $10.0$ , et les graphes recourent à une échelle des abscisses logarithmique. La figure 6.28 présente l'évolution de la distance moyenne des agents au barycentre pour différentes valeurs des liens sociaux, tandis que la figure 6.29 illustre l'évolution de la distance moyenne parcourue par les agents et le barycentre pour les mêmes paramètres.

On constate plusieurs résultats intéressants :

- on retrouve l'opposition très nette entre des liens positifs, qui entraînent une proximité des deux agents, et des liens négatifs qui conduisent au contraire à une répulsion entre les deux agents et une distance au barycentre importante ;
- plus la variance sociale est faible, plus les distances parcourues en moyenne par le barycentre ou les deux agents sont faibles. Celles-ci tendent d'ailleurs vers 0 lorsque la variance tend elle-même vers 0. Une absence de liberté vis à vis de la norme sociale conduit à un gel des évolutions du système, tout comme une trop faible variance naturelle empêchait un système d'évoluer ;
- plus la variance est faible, plus la distance au barycentre est faible dans le cas de liens sociaux positifs, et plus elle est importante et proche de la limite dans le cas de liens répulsifs. Lorsque la variance tend vers 0, la distance tend vers 0 quelle que soit la force des liens positifs, et vers 0.70 quelle que soit la valeur des liens négatifs. Pour des valeurs plus importantes de la variance, la liberté engendrée par une plus faible adéquation à la norme sociale se traduit plus ou moins selon la force des liens sociaux, et nous retrouvons les résultats des paragraphes précédents (la figure 6.24 est ainsi une vue en coupe de la figure 6.28 pour une valeur de la variance égale à 0.001 ou 0.00001). Pour une variance

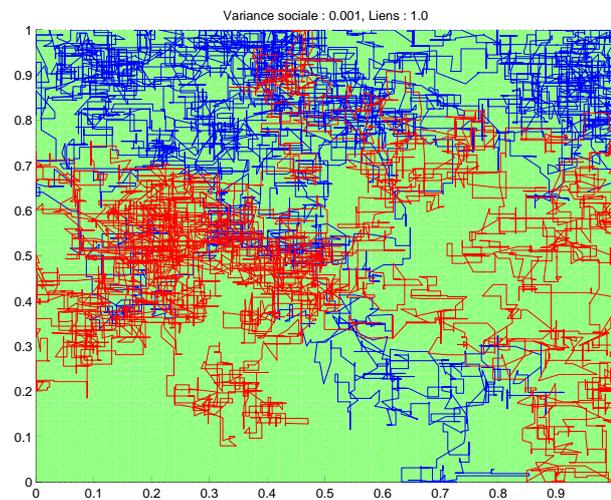


FIG. 6.26 – Evolution de deux locuteurs unis par deux liens positifs symétriques de valeur 1.0 pour une variance sociale égale à 0.001

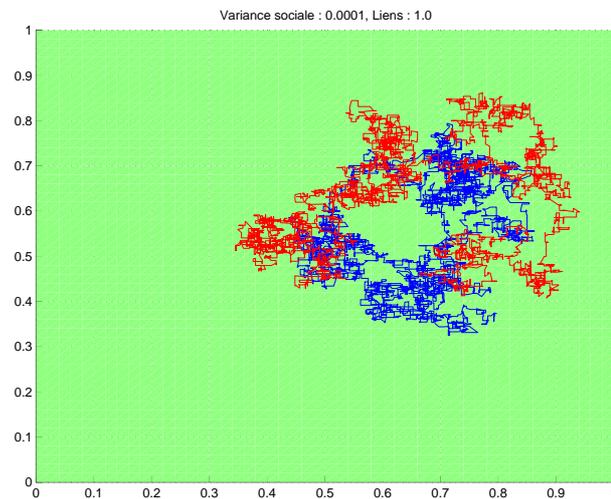


FIG. 6.27 – Evolution de deux locuteurs unis par deux liens positifs symétriques de valeur 1.0 pour une variance sociale égale à 0.0001

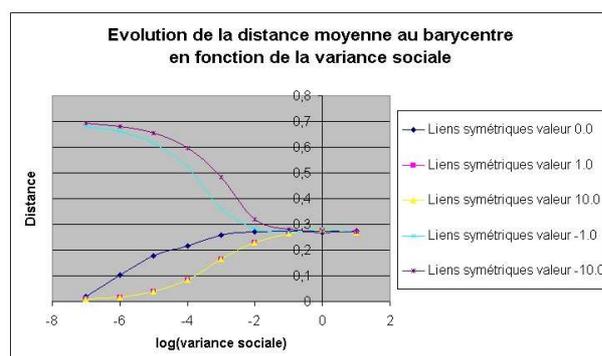


FIG. 6.28 – Evolution de l'entropie du système et de la distance moyenne parcourue par les locuteurs en fonction de la variance de la distribution sociale

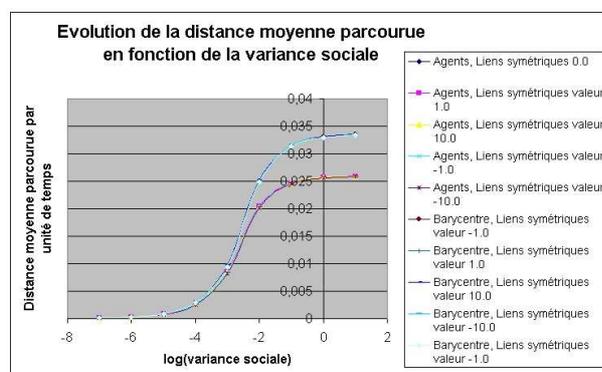


FIG. 6.29 – Evolution de la distance moyenne parcourue par les agents en fonction de la variance sociale

sociale supérieure ou égale à 0.1, les courbes se confondent, et la valeur des liens sociaux ne joue plus aucun rôle. La liberté laissée aux agents pour leur évolution vis à vis de la norme sociale est telle que leurs liens deviennent non pertinents ;

- La courbe de la distance moyenne au barycentre tend vers 0 même lorsque les liens sociaux entre agents sont nuls ; ceci peut s'expliquer par le fait qu'une variance de plus en plus faible permet de moins en moins de déviations autour de la norme sociale. Plus la variance est faible, moins les agents ont donc tendance à bouger quand leurs liens sont nuls, puisque la déviation sociale prend toujours une valeur nulle. Comme en outre la situation initiale des simulations correspond à deux agents situés tout deux au centre de l'espace des paramètres, la distance de ceux-ci au barycentre est initialement nulle et demeure très faible si ceux-ci ne peuvent que très peu se déplacer. En fait, la courbe théorique pour une durée de temps infinie, une variance non nulle et des liens nuls devrait être horizontale, mais les valeurs des indicateurs sont établies durant une période de temps finie, ce qui conduit à la courbe observée ;

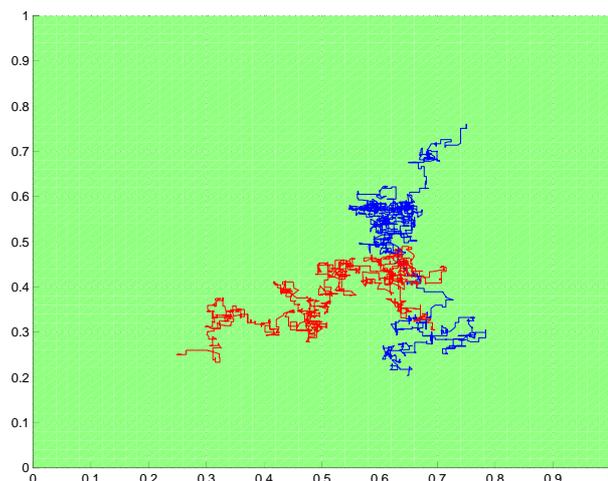


FIG. 6.30 – Evolution de deux locuteurs unis par des liens positifs symétriques et initialement localisés en deux positions distinctes de l’espace des possibles

### Contact entre deux locuteurs de systèmes différents

Une dernière expérience très simple avant d’examiner des communautés de plus de deux agents consiste à observer le déroulement d’une situation de “contact”, où deux agents unis par des liens positifs et initialement localisés en deux positions distinctes de l’espace des possibles sont mis en contact. Dans cette expérience, nous allons envisager le cas de liens symétriques et celui de liens asymétriques.

La figure 6.30 illustre la situation où deux agents unis par deux liens symétriques de valeur 10.0 sont mis en contact, alors que la figure 6.31 présente le cas où l’un des deux liens prend une valeur de 10.0 et le second est nul.

Dans le premier cas, on observe que les agents convergent rapidement l’un vers l’autre et commencent alors à évoluer de façon stochastique selon les marges de liberté laissées par le choix de la variance sociale. Les deux agents adoptent la même démarche pour se rejoindre.

Par contraste, dans le second cas, où c’est l’agent de couleur bleue qui est fortement influencé par celui de couleur rouge, on observe bien un rapprochement des agents, mais uniquement à l’initiative de l’agent bleu, qui rejoint un agent rouge qui se déplace de façon totalement aléatoire (puisque sa déviation sociale à chaque pas de temps est toujours nulle). Ceci est tout à fait en accord avec la structure de la norme sociale en vigueur, qui laisse l’agent rouge libre et contraint l’agent bleu à rester proche du premier.

### Influence du nombre d’individus

Après avoir étudié en détail l’influence de différents paramètres sur l’évolution de deux agents, il est temps d’examiner l’impact du nombre d’agents sur les évolutions du système. En effet, si l’on se remémore les expériences de Nettle et le problème de la norme communautaire (voir chapitre 2), nos prédictions sont que le système devrait évoluer moins vite lorsque le nombre d’agents augmente.

Les figures 6.32 et 6.33 permettent de confirmer cette intuition. Comme pour les expériences précédentes, la première courbe représente la distance moyenne au barycentre, et la seconde les

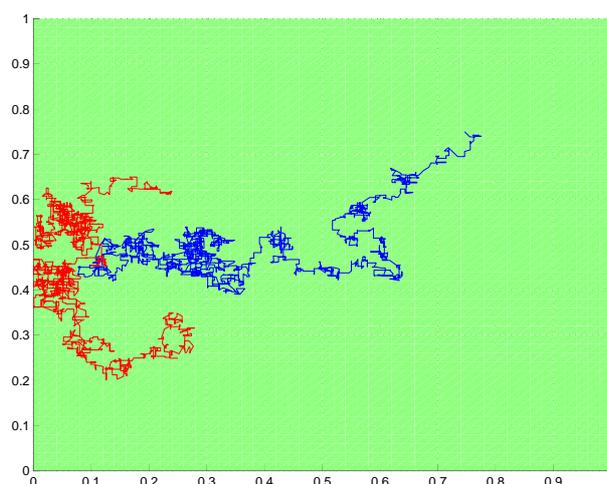


FIG. 6.31 – Evolution de deux locuteurs unis par des liens positifs ou nuls asymétriques et initialement localisés en deux positions distinctes de l'espace des possibles

distance moyennes parcourues par pas de temps par le barycentre des agents et ces derniers.

L'examen de la figure 6.32 permet d'observer une croissance lente et limitée par une asymptote horizontale des courbes qui représentent des liens symétriques négatives ou nuls : lorsque le nombre d'agents augmente, la distance entre les agents augmente légèrement et de plus en plus lentement pour atteindre une valeur limite, d'autant plus importante que la force des liens répulsifs est importante.

Au contraire, la distance des agents au barycentre diminue avec le nombre d'agents dans les cas de liens positifs, mais là encore, une asymptote supérieure à 0 semble limiter la décroissance. Cette asymptote est d'autant plus faible que la valeur des liens est grande.

Pour le cas de liens positifs, la diminution de la distance moyenne au barycentre avec le nombre d'agents peut s'expliquer simplement par la plus grande force d'attraction qu'exerce alors la communauté sur un de ses membres. La déviation sociale est fortement marquée en direction du barycentre des agents, et un agent a donc de plus en plus de mal à s'éloigner de celui-ci au fur et à mesure que le nombre d'agents augmente. Néanmoins, une asymptote horizontale non nulle vient limiter la décroissance de la distance moyenne au barycentre. Celle-ci est vraisemblablement due à la variance sociale non nulle qui empêche les agents de se rejoindre tout à fait.

Une analogie peut permettre de mieux comprendre la situation : lorsqu'une planète a consommé toutes ses réserves de combustible nucléaire, après un passage par le stade de géante rouge, la pression exercée par les réaction nucléaires ne parvient plus à compenser l'attraction gravitationnelle, et l'étoile s'effondre sur elle-même. Si la masse initiale de cette étoile est inférieure à 1.44 masse solaire, l'effondrement est stoppé au bout d'un moment car une pression de dégénérescence vient équilibrer les forces gravitationnelles. Cette pression est due aux répulsions entre les électrons des atomes, selon le principe quantique d'exclusion de Pauli (on parle aussi de pression de Fermi). On observe alors une naine blanche extrêmement dense.

Si l'étoile a une masse comprise entre 1.44 et 3.2 masses solaires, la pression créée par les électrons dégénérés n'est plus suffisante pour compenser la gravité, et un nouvel effondrement

se produit, au cours duquel les électrons pénètrent le noyau des atomes et fusionnent avec les protons. Ce procédé de neutronisation produit un grand nombre de neutrons, et les interactions fortes entre eux produisent une nouvelle force de pression qui parvient à stopper l'effondrement gravitationnel. Une étoile à neutrons est née.

De la même façon que les pressions de dégénérescence viennent empêcher une étoile de se transformer en trou noir, une variance sociale non nulle peut équilibrer les forces d'attraction entre les agents lorsque le groupe devient trop compact, et conduire à une valeur minimale non nulle de la distance moyenne au barycentre. De la même façon qu'une masse plus importante peut conduire à une étoile à neutrons plutôt qu'à une naine blanche, des liens attractifs de plus forte valeur conduiront à un équilibre caractérisé par une plus faible distance moyenne au barycentre.

Il nous paraît plus délicat d'expliquer l'accroissement de la distance moyenne entre agents dans le cas de liens négatifs ou nuls. Dans le cas de liens nuls, les agents se déplacent aléatoirement sur la carte. Plus le nombre d'agents sera important, plus ceux-ci se répartiront dans l'espace des paramètres, et plus le barycentre sera proche de la position centrale de l'espace. Nous pensons que cette position centrale a pour conséquence de réduire la distance moyenne entre les agents et le barycentre, puisqu'intuitivement, les distances au barycentres seraient plus importantes si celui-ci était placé dans un des angles de l'espace des paramètres. L'effet est néanmoins peu marqué, ce qui explique la faible croissance des valeurs avec le nombre d'agents.

Pour le cas de liens négatifs, un examen détaillé des trajectoires évolutives des agents permet de voir une distribution de ces agents le long des frontières de l'espace des paramètres. Un accroissement du nombre d'agents rend plus difficile toute progression d'un agent vers le barycentre situé au centre de l'espace, puisque les agents plus nombreux situés sur l'arête opposée de l'espace des paramètres repoussent l'agent sur sa propre arête. Ceci conduit à un accroissement modéré de la distance moyenne au barycentre.

Si l'on s'intéresse maintenant à l'évolution des distances moyennes parcourues, on constate que celle du barycentre du système décroît de façon approximativement exponentielle pour se diriger vers 0 pour un grand nombre d'agents, et ce de façon très similaire quelle que soit la valeur des liens tissés entre les agents. On constate toutefois un faible écart entre les courbes, et un ordonnancement qui les classe selon la force des liens pour les cas qu'elles représentent : les distances les plus faibles correspondent aux liens les plus négatifs, puis viennent les liens faiblement négatifs. . . Les courbes pour les liens positifs sont cependant pratiquement superposées sur la figure.

La décroissance de la distance moyenne parcourue par le barycentre s'explique dans tous les cas par le fait qu'un plus grand nombre d'agents minimise les variations du barycentre par le simple phénomène de moyennage. En effet, les mouvements d'un agent seront d'autant moins reflétés au niveau du barycentre qu'un grand nombre d'agents sera pris en compte dans le calcul de ce barycentre.

L'évolution des distances moyennes parcourues par les agents est qualitativement différente de celle des barycentres. Pour des liens nuls, on observe que la distance moyenne parcourue reste constante quelque soit le nombre d'agents. Ceci est bien sûr normal puisque les agents ne s'influencent pas les uns les autres et que le nombre d'agents ne joue donc aucun rôle. Pour des liens positifs, on observe une légère augmentation de la distance moyenne parcourue par les agents avec le nombre d'agents. Cette augmentation peut s'expliquer par la plus grande cohérence d'une population lorsque le nombre d'agents augmente. Encore une fois, une plus forte norme sociale

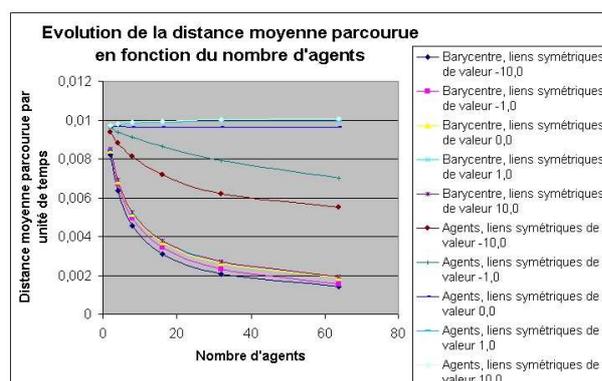


FIG. 6.32 – Evolution de la distance moyenne parcourue par le barycentre et les agents en fonction du nombre d'agents

permet aux agents d'évoluer de façon plus congruente, et augmente la vitesse de leurs déplacements. Enfin, dans le cas de liens négatifs, la distance moyenne parcourue diminue avec le nombre d'agents. Ceci peut s'expliquer par le fait qu'un plus grand nombre d'agents maintient plus fortement ceux-ci sur les bords de l'espace des paramètres, et rend en conséquence leur évolution plus difficile.

L'écart entre les distances moyennes parcourues par les agents ou le barycentre selon la valeur des liens n'existe pas pour les situations à deux agents. Cette observation rejoint les résultats de la première expérience de cette série, où l'on observait une indépendance de la distance moyenne parcourue par le barycentre ou les agents vis à vis des liens sociaux.

Pour plus de deux agents cette fois, on peut conclure à un accroissement du rythme d'évolution du barycentre et des agents lorsque la valeur des liens sociaux augmente (depuis des valeurs très négatives jusqu'à des valeurs très positives). Cette effet est marqué pour des valeurs négatives différentes, et entre des valeurs négatives et positives, mais beaucoup moins pour des valeurs positives différentes. En particulier, la distinction entre liens positifs et négatifs est particulièrement importante en ce qui concerne les rythmes d'évolution des agents lorsque leur nombre augmente.

Les résultats principaux à retenir de cette expérience est (i) qu'un accroissement du nombre d'agents diminue le rythme d'évolution d'un système et son exploration de l'espace des possibles et (ii) que cet accroissement diminue la diversité du système en cas de liens positifs, et l'augmente en cas de liens négatifs. Une diversité minimale subsiste cependant quelque soit le nombre d'agents si les liens sont positifs, et la diversité est aussi limitée dans le cas de liens négatifs. Enfin, (iii) le rythme d'évolution des agents unis par des liens positifs s'accroît légèrement avec une augmentation du nombre d'agents, alors qu'il diminue fortement si les liens sont négatifs.

### Plage de distribution des liens sociaux

Pour conclure cette deuxième série d'expériences centrées sur les contraintes sociales, nous pouvons étudier rapidement comment la distribution géographique d'un réseau social peut influencer les caractéristiques de l'évolution d'un macro-système linguistique.

Plus concrètement, nous considérons 10 agents disposés en anneau. Cette topologie mime un espace géographique composé d'une seule dimension, ce qui peut paraître très éloigné de la

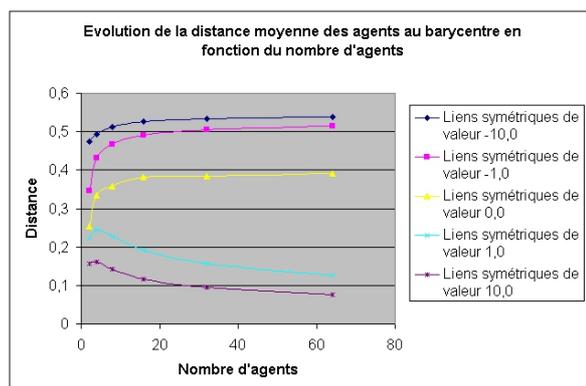


FIG. 6.33 – Evolution de la distance moyenne des agents au barycentre du système en fonction du nombre d’agents

situation d’un espace bidimensionnel plus classique. Toutefois, nous pensons que les résultats établis avec cette topologie monodimensionnelle peuvent être étendus sans problème à une topologie en grille bidimensionnelle.

Chaque agent peut établir des liens avec les agents situés à sa gauche et à sa droite. Nous souhaitons comparer trois schémas de connexion des liens, qui reflètent une distribution plus ou moins “condensé” des liens d’un agent vers les agents qui l’entourent. Dans le premier cas, l’agent n’établit des liens (de valeur 10.0) qu’avec les deux agents qui l’entourent immédiatement, comme le montre la table 6.1. Ceci correspond à une situation où un agent n’établit des relations qu’avec ses plus proches voisins géographiques. Dans le second cas, illustré par la table 6.2, l’agent établit des liens de valeur 6.0 avec les agents immédiatement à sa droite et à sa gauche, ainsi que des liens de valeur 4.0 avec les agents situés 2 rangs à sa gauche et à sa droite. Dans ce cas de figure, les liens tissés par un agent avec d’autres restent limités dans leur extension géographique, mais s’étendent toutefois sur une plus longue distance que dans le premier cas.

Enfin, la troisième et dernière situation correspond à une distribution “étalée”, où chaque agent établit des liens de valeur 2.0 avec l’ensemble des agents de l’anneaux (table 6.3). Dans ce dernier cas, un agent communique avec l’ensemble des agents de l’environnement, mais selon des liens assez faibles.

On constate que pour les trois distributions, la somme des valeurs des liens pour chaque agent est toujours égale à 20.0, et symétrique de part et d’autre des agents. Ceci permet de comparer des situations où la force globale des liens qui lient un agent au reste de la communauté est identique.

L’analyse des données des simulations conduit aux résultats suivants :

- la distance moyenne parcourue par les agents est constante quelle que soit la distribution adoptée ;
- la distance moyenne parcourue par le barycentre du système est également constante ;
- la distance moyenne des agents au barycentre évolue pour les trois distributions comme illustré par la figure 6.34. On constate une décroissance de la distance des agents au barycentre avec l’étalement des liens le long de la dimension géographique.

---

0	10	0	0	0	0	0	0	0	10
10	0	10	0	0	0	0	0	0	0
0	10	0	10	0	0	0	0	0	0
0	0	10	0	10	0	0	0	0	0
0	0	0	10	0	10	0	0	0	0
0	0	0	0	10	0	10	0	0	0
0	0	0	0	0	10	0	10	0	0
0	0	0	0	0	0	10	0	10	0
0	0	0	0	0	0	0	10	0	10
10	0	0	0	0	0	0	0	10	0

TAB. 6.1 – Distribution géographique condensée des liens pour les agents en anneau

0	6	4	0	0	0	0	0	4	6
6	0	6	4	0	0	0	0	0	4
4	6	0	6	4	0	0	0	0	0
0	4	6	0	6	4	0	0	0	0
0	0	4	6	0	6	4	0	0	0
0	0	0	4	6	0	6	4	0	0
0	0	0	0	4	6	0	6	4	0
0	0	0	0	0	4	6	0	6	4
4	0	0	0	0	0	4	6	0	6
6	4	0	0	0	0	0	0	6	0

TAB. 6.2 – Distribution géographique intermédiaire des liens pour les agents en anneau

Nous pouvons donc conclure de cette dernière expérience que pour un nombre d'agents fixé et pour une projection à plus ou moins grande distance des liens sociaux dans l'espace géographique, la vitesse d'évolution des systèmes linguistiques ou du barycentre de ces systèmes reste constante. Le seul impact est une plus grande proximité entre les agents et leur barycentre, c'est à dire une plus faible diversité linguistique du macro-système, lorsque le réseau social est largement distribué spatialement.

### 6.2.3 Troisième série d'expériences : Evolution d'un système linguistique communal sous contraintes internes et sociales

Au cours des deux premières séries d'expériences, nous avons étudiés respectivement le cas d'un seul système soumis à des contraintes naturelles, puis celui de plusieurs systèmes en interaction, mais en l'absence de contraintes naturelles et distribuées (seules des contraintes sociales étaient mises en œuvre). Pour conclure notre étude générale, il nous reste à croiser les deux types de contraintes.

#### Problème de la détermination d'indicateurs d'évolution du système

Un des problèmes qui apparaît lorsque l'on s'intéresse à l'évolution de systèmes d'agents en interaction dans un paysage énergétique est la difficulté de définir des indicateurs d'évolution pertinents.

2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2

TAB. 6.3 – Distribution géographique étalée des liens pour les agents en anneau

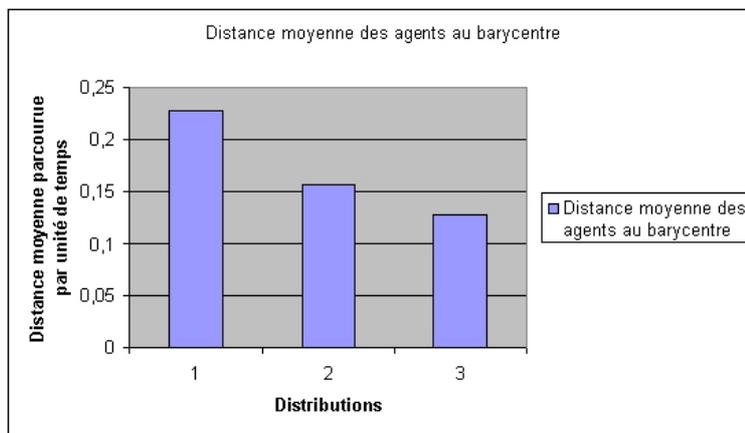


FIG. 6.34 – Evolution de la distance moyenne parcourue par le barycentre pour les trois situations

Nous pouvons bien sûr toujours recourir aux indicateurs définis et utilisés dans la seconde série d’expériences, à savoir la distance moyenne au barycentre des agents, et la distance moyenne parcourue par les agents et le barycentre de ceux-ci au cours du temps. Toutefois, ces indicateurs ne nous permettront pas de détecter les changements d’état à la fois des agents et du barycentre.

Il est également possible pour les agents de recourir à l’indicateur défini lors de la première série d’expériences, à savoir le nombre de transitions d’un attracteur à un autre. Il serait tentant d’appliquer également cet indicateur au barycentre, mais le processus de moyennage à l’origine du calcul de cet élément fait qu’il ne suit pas du tout le relief du paysage. Nous nous contenterons donc de mesurer comme dans la deuxième série d’expérience la distance moyenne parcourue par le barycentre, tout en gardant à l’esprit que cet indicateur n’est pas aussi pertinent que précédemment.

Afin de pouvoir étudier la façon dont les agents d’un système se comportent les uns vis à vis des autres dans le paysage énergétique, nous devons être en mesure de détecter des phases où les agents n’occupent pas le même bassin d’attraction.

Puisque notre but ici est d’évaluer les conséquences des interactions sociales sur les évolutions de systèmes linguistiques dans un paysage énergétique, et non pas d’étudier ce dernier, nous

pouvons mettre à profit notre connaissance de celui-ci pour définir des indicateurs pertinents.

Nous allons ainsi utiliser un indicateur proche de l'entropie que nous avons définie plus haut. Nous avons souligné alors les difficultés d'appliquer cette mesure à des systèmes composés d'un petit nombre d'agents, et ce problème doit être résolu lors de la définition de notre nouvel indicateur.

Nous travaillerons par la suite avec le même paysage énergétique, composé de 4 puits identiques, que nous avons utilisé dans les premières expériences de la première série. Etant donné un agent, il est facile de déterminer le bassin d'attraction dans lequel il se trouve connaissant sa position.

Pour chacun des 4 bassins d'attraction, nous pouvons faire la somme à chaque pas de temps des agents qui se trouvent en son sein, et diviser cette somme par le nombre d'agents total. Nous définissons alors le nouvel indicateur entropique de la façon suivante :

$$entropy(t) = - \sum_{i=1}^4 p_i \ln(p_i)$$

où  $p_i$  est le ratio du nombre d'agents qui appartiennent au  $i$ ème bassin d'attraction sur le nombre total d'agents.

La mesure précédente prendra sa plus faible valeur lorsque tous les agents du système seront situés dans le même bassin d'attraction. Au contraire, la valeur la plus importante sera atteinte lorsque les agents seront les plus distribués possibles dans les différents bassins. Le problème de l'ancienne définition de l'entropie est corrigé.

Connaître l'évolution de l'entropie au cours du temps permettra d'avoir une idée sur les changements au sein du système multi-agents, tandis que la valeur moyenne de l'entropie pour une simulation renseignera sur la répartition moyennes des agents dans les différents bassins d'attraction, et donc sur la *diversité* du macro-système.

### **Etude de l'évolution de deux agents selon différentes configurations sociales**

Nous allons donc commencer comme dans la deuxième série d'expériences par étudier l'évolution de deux agents en fonction de la valeur des liens symétriques qui les unissent.

Les figures 6.35 et 6.36 donnent respectivement la distance moyenne des agents au barycentre et la distance moyenne parcourue par les agents et le barycentre du système.

Les courbes observées sont en tout point similaires aux courbes que nous avons observées pour la situation en absence de contraintes naturelles. L'ensemble des analyses faites plus haut peut donc être appliqué ici, et il semble que pour cette expérience, l'existence de contraintes naturelles ne perturbent pas les caractéristiques qualitatives de l'évolution du système.

Pour confirmer cet état de fait, la figure 6.37 représente l'entropie moyenne du système au cours d'une simulation. La courbe observée s'explique assez facilement : la valeur de l'entropie importante et constante pour les valeurs de liens négatives s'explique par le fait que les répulsions entre les deux agents empêchent ceux-ci de se trouver dans le même bassin d'attraction. L'entropie vaut donc (presque) toujours la valeur maximale prise lorsque les agents sont dans deux bassins d'attraction différents, soit  $2.0 \times 0.5 \times \log(0.5) = \log(2) = 0.69$ . Ceci est vrai quelle que soit la valeur de la variance sociale.

Pour des valeurs des liens positives, on observe une diminution de l'entropie, qui caractérise le fait que les agents se trouveront plus souvent dans le même bassin d'attraction. On remarque

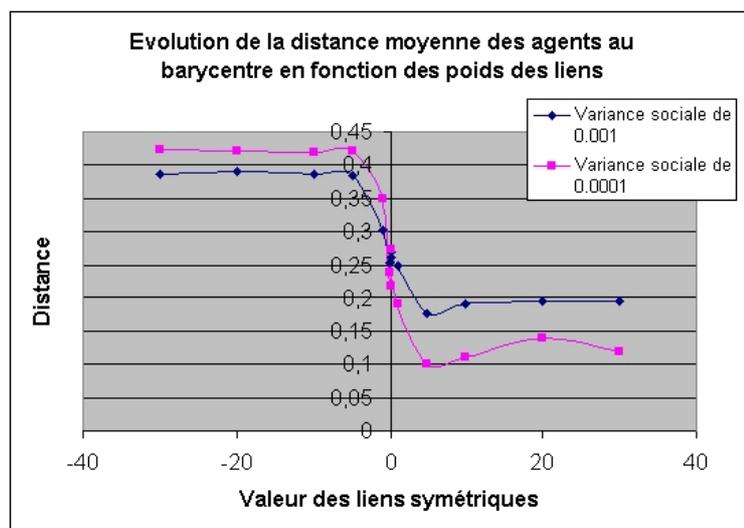


FIG. 6.35 – Evolution de la distance moyenne au barycentre en fonction des liens (symétriques) entre agents en présence de contraintes naturelles

également le fait que la courbe pour la variance sociale la plus faible se situe sous la courbe de variance sociale plus importante, toujours pour des liens positifs. Ceci est en accord avec la plus grande proximité des deux agents dans le second cas, et donc avec le fait qu'ils auront très peu tendance à se trouver dans des bassins d'attraction différents.

Entropie et distance des agents au barycentre sont donc globalement congruentes pour cette expérience.

### Un exemple d'évolution de deux agents unis par un lien positif

Nous pouvons illustrer l'étude précédente par les trajectoires évolutives d'une simulation particulière. Dans cette expérience, les deux agents sont reliés par des liens positifs de valeur 1.0.

Figure 6.38, l'observation de la situation après 5,000 pas de temps permet de constater que l'agent bleu a quitté le bassin d'attraction qu'il occupait avec le second agent rouge, et a évolué rapidement en franchissant deux barrières d'énergie successives. Les deux agents se trouvent très éloignés l'un de l'autre, et cette situation ne répond pas de façon adéquate aux contraintes sociales. On s'attend donc logiquement à une réaction du système pour corriger ce déséquilibre.

La situation après 7,000 pas de temps permet de constater que le système a réagi pour corriger partiellement le déséquilibre précédent (figure 6.39). Pour cela, l'agent rouge a franchi à son tour une barrière énergétique sur les traces de l'agent bleu, et se trouve maintenant dans le bassin d'attraction adjacent au sien. Cette situation n'est pas encore optimale, puisque les deux agents ne se trouvent pas dans le même bassin d'attraction.

Après 8,000 pas de temps (figure 6.40, les deux agents se trouvent à nouveau dans le même bassin d'attraction et le système satisfait à la fois aux contraintes naturelles et sociales. Ce n'est pas l'agent rouge qui a rejoint l'agent bleu, mais au contraire celui-ci qui a rebroussé chemin pour gagner le bassin d'attraction occupé par l'agent rouge et qu'il avait traversé auparavant.

Nous pouvons rapprocher cet exemple de comportement de phénomènes linguistiques réels,

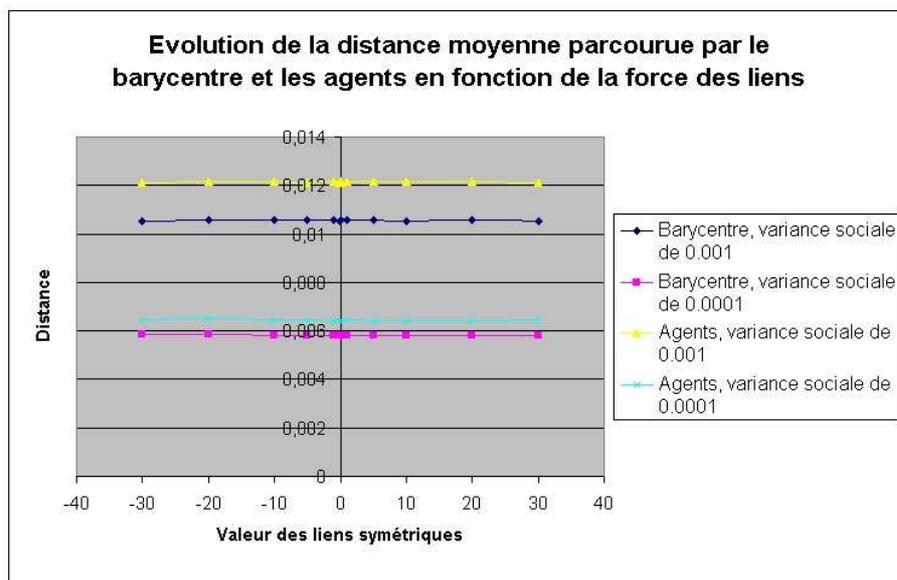


FIG. 6.36 – Evolution de la distance moyenne parcourue par les agents et le barycentre en fonction des liens (symétriques) entre agents en présence de contraintes naturelles

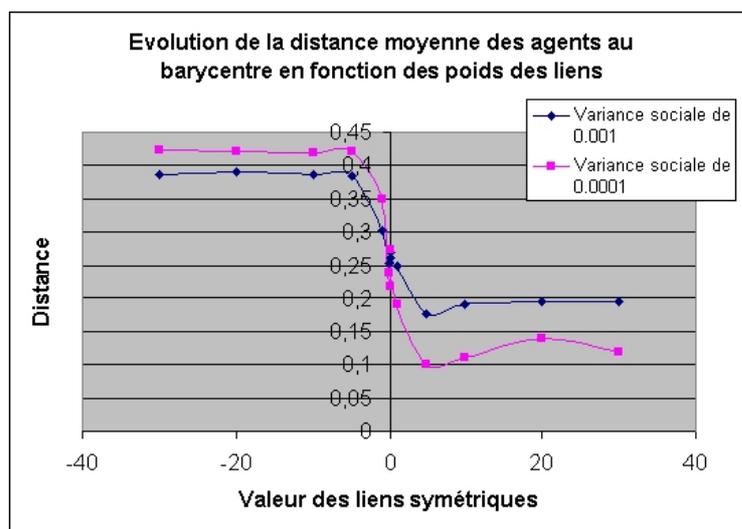


FIG. 6.37 – Evolution de l'entropie du système en fonction des liens (symétriques) entre agents en présence de contraintes naturelles

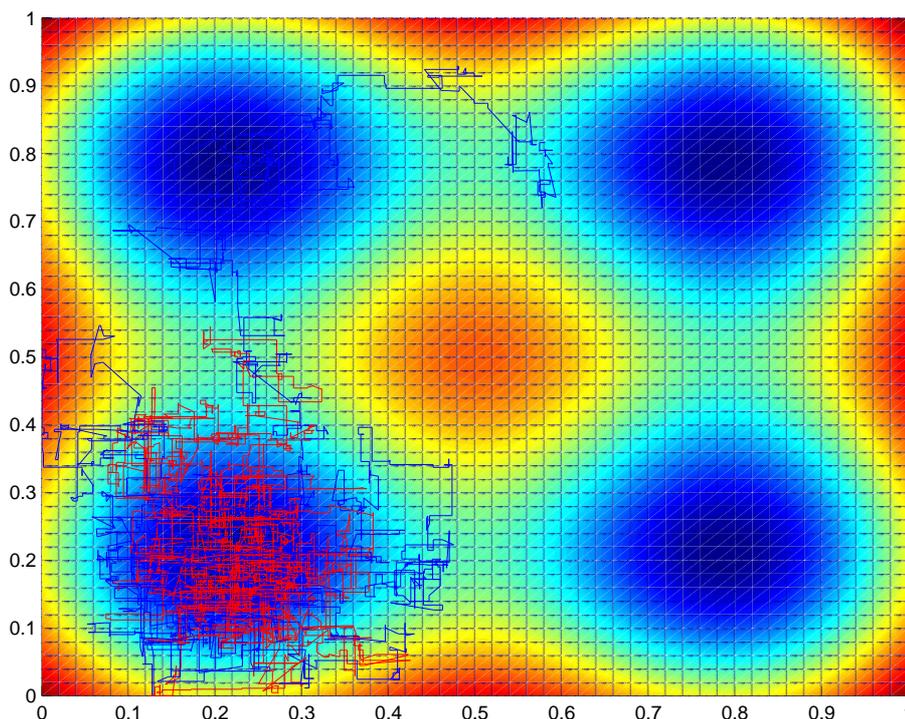


FIG. 6.38 – Trajectoires évolutives de deux agents liés par des liens positifs de valeur 1.0, 5000 pas de temps

comme celui que nous avons décrit au chapitre 2, et qui portait sur la fusion puis la fusion de deux phonèmes. Si l'on assimile chaque agent à une couche de la société, on peut observer pendant un certain temps une différenciation entre les deux strates sociales, puis un rapprochement et l'adoption d'une norme linguistique commune. Des phénomènes contingents sont à l'origine de ces phénomènes transitoires, qui nous semblent se produire dans la réalité.

### **Impact de la variance sociale sur les évolutions d'un système composé de deux agents**

De la même façon que pour l'étude des liens, l'étude de l'évolution des systèmes en fonction de la variance sociale conduit globalement aux mêmes résultats que dans la seconde série d'expériences. Nous ne nous attardons donc pas à reproduire ici les résultats, et passons à un cas plus intéressant.

### **Etude de systèmes comportant différents nombres d'agents**

Toujours en suivant le décours des expériences de la deuxième série, nous allons enfin examiner l'influence du nombre d'agents sur l'évolution d'un macro-système linguistique

Les figures 6.41 6.42 et 6.43 donnent dans l'ordre la distance moyenne des agents au barycentre, la distance moyenne parcourue par le barycentre, et l'entropie du système.

Comme dans le cas d'un paysage énergétique plat, la distance parcourue en moyenne par le barycentre diminue avec le nombre d'agents. Les mêmes raisons que celles invoquées précédem-

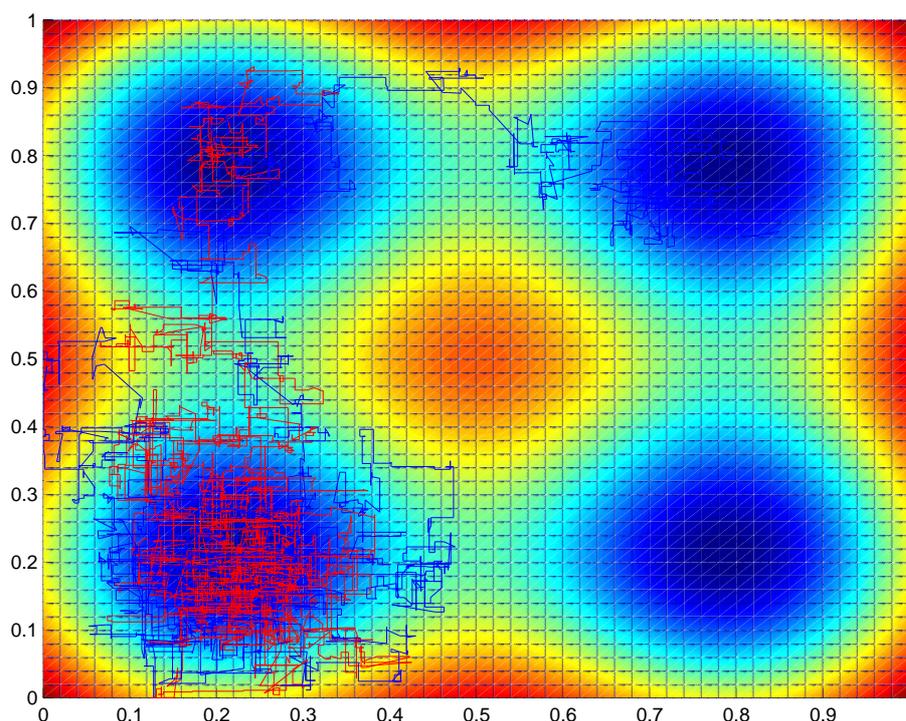


FIG. 6.39 – Trajectoires évolutives de deux agents liés par des liens positifs de valeur 1.0, 7000 pas de temps

ment peuvent également être appliquées ici, modulo le fait que les agents ont tendance à suivre les reliefs du paysage énergétique.

Les valeurs de l'entropie du macro-système et de la distance moyenne des agents au barycentre sont de nouveau congruentes, quelle que soit la valeur des liens sociaux ou le nombre d'agents.

En outre, les schémas d'évolution de ces deux grandeurs en fonction du nombre d'agents sont similaires à celui de la distance des agents au barycentre pour l'expérience correspondante de la deuxième série avec un paysage énergétique plat, y compris au niveau d'une légère croissance des courbes pour les liens positifs avant une décroissance marquée. Nous pouvons donc à nouveau translater les explications de l'expérience de la deuxième série à celle-ci :

- en ce qui concerne l'évolution de la distance des agents au barycentre, nous pouvons constater que celle-ci augmente légèrement pour les valeurs de liens négatives, ce qui est dû encore une fois aux répulsions entre systèmes qui auront tendance, tout en restant dans les bassins d'attraction, à se déplacer vers les frontières de l'espace des paramètres au fur et mesure de l'augmentation du nombre d'agents ;
- l'entropie du système croît pour les valeurs de liens négatives, ce qui est en accord avec l'idée qu'un plus grand nombre d'agents opposés à cause de liens répulsifs se répartiront de la façon la plus distribuée possible dans les bassins d'attraction, et augmenteront donc d'autant la valeur de l'entropie ;
- pour des liens nuls, la simple augmentation du nombre d'agents conduira à une distribution plus homogène des agents dans les bassins d'attraction (diminution des effets de "petite taille" du système), et conduira donc aussi à une légère augmentation de l'entropie. On

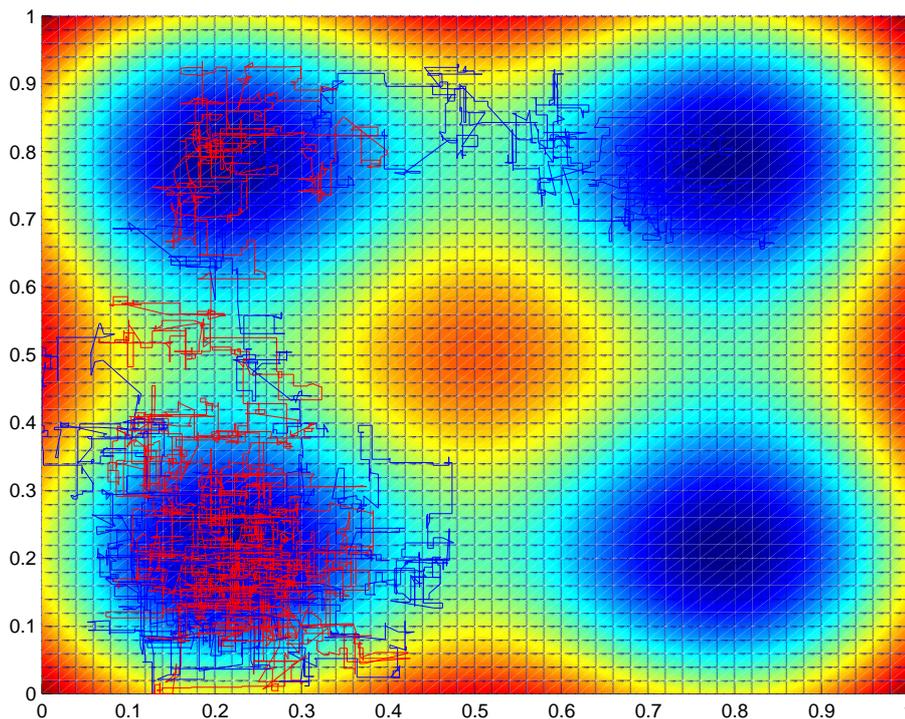


FIG. 6.40 – Trajectoires évolutives de deux agents liés par des liens positifs de valeur 1.0, 8000 pas de temps

constate d'ailleurs que la courbe pour les liens nuls se trouve sous les courbes pour les liens négatifs.

- pour les liens de valeur 1.0, on constate une légère croissance de l'entropie et de la distance moyenne des agents au barycentre, avant une décroissance marquée. L'accroissement de l'entropie lors de l'augmentation de la taille du système de 2 à 4 puis 8 agents nous semble correspondre à la même logique que celui de la courbe de liens nuls. La simple augmentation du nombre d'agents et la relative faiblesse des liens conduit à une meilleure occupation des bassins d'attraction, et donc à une plus grande entropie. Cependant, quand l'augmentation des agents se fait trop importante, la grande cohérence du système vient limiter les changements d'états pour les agents (à cause d'une déviation sociale marquée en faveur de la communauté), qui se retrouvent à être beaucoup plus nombreux dans les mêmes bassins d'attraction, ce qui diminue la valeur de l'entropie ;
- ce dernier schéma explicatif se confirme dans l'évolution de la courbe pour des liens de valeur 10.0. L'épisode de croissance de cette courbe est beaucoup plus restreint que celui de la précédente, ce qui s'explique par la plus grande cohésion du groupe. Cette cohésion contrebalance beaucoup plus vite l'augmentation d'entropie liée à l'accroissement du nombre d'agents, et la diminution d'entropie se produit beaucoup plus rapidement.

Pour plus de lisibilité, nous avons représenté l'évolution de la distance moyenne parcourue par les agents en fonction du nombre de ceux-ci sur un graphe indépendant. La figure 6.44 permet de constater un certain entrelacement des courbes pour les différentes valeurs de liens, mais également une décroissance peu marquée pour une valeur des liens de -1.0, et une décroissance cette fois plus nette pour une valeur de -10.0. Ceci est encore une fois similaire à la situation

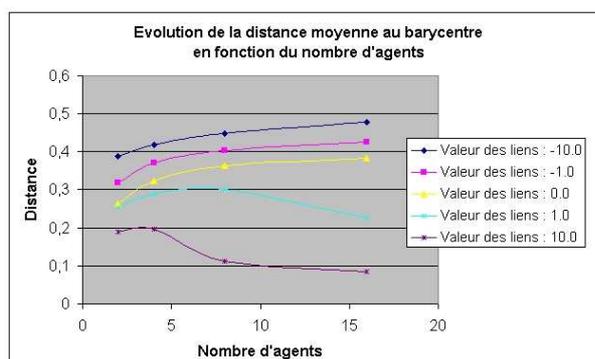


FIG. 6.41 – Evolution de la distance moyenne des agents au barycentre du système en fonction du nombre d'agents (liens symétriques de valeur 1.0 entre tous les agents)

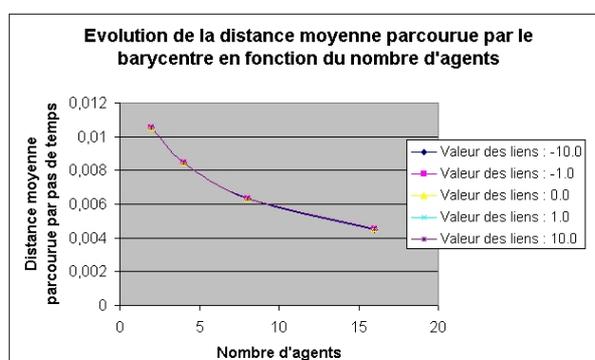


FIG. 6.42 – Evolution de la distance moyenne parcourue par le barycentre en fonction du nombre d'agents

observée pour un paysage énergétique plat. Néanmoins, on peut noter que si les tendances sont les mêmes que pour la situation en absence de contraintes naturelles et distribuées, les variations sont beaucoup plus faibles, de l'ordre de quelques pourcents au plus. La prise en compte de contraintes naturelles et distribuées vient donc minimiser les variations de la distance moyenne parcourue par les agents en fonction de leur nombre.

### Conclusions préliminaires

Il apparaît au travers des expériences précédentes et en première approche que, pour des valeurs intermédiaires des variances naturelle et sociale, les contraintes internes ne modifient pas les résultats qualitatifs qui avaient été obtenus avec un paysage énergétique plat, mais imposent seulement des contraintes sur la façon dont les agents se déplacent dans l'espace des paramètres. Une bonne congruence est observée en outre entre la valeur de l'entropie d'un macro-système et la distance moyenne des agents de ce système à leur barycentre. Ce dernier indicateur, que nous avons assimilé à une mesure de diversité linguistique de façon légère lors de la deuxième série d'expériences, se révèle en fait valide. A l'inverse, l'existence de contraintes sociales ne vient pas modifier la façon dont les agents se déplacent dans le paysage, c'est à dire selon une dynamique d'équilibres ponctuels.

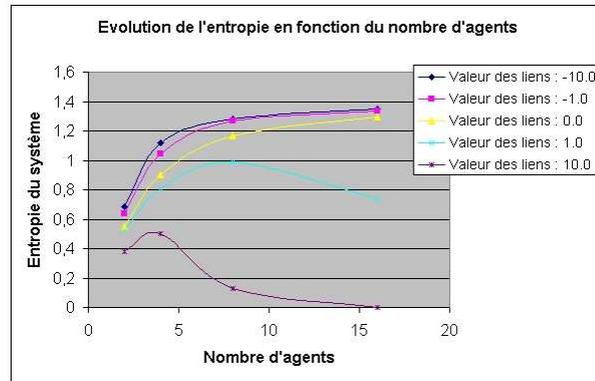


FIG. 6.43 – Evolution de l'entropie du système en fonction du nombre d'agents

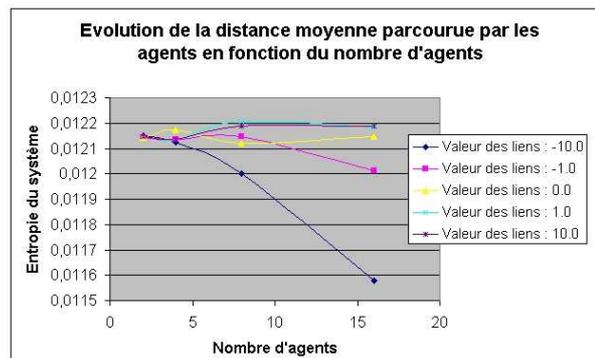


FIG. 6.44 – Evolution de la distance moyenne parcourue par les agents en fonction du nombre d'agents

Même si des expériences supplémentaires sont à envisager pour bien confirmer ce fait, la relative indépendance entre les contraintes internes et les contraintes sociales va apparemment dans le sens de l'*existence de domaines opérationnellement distincts au sens de Maturana : les contraintes sociales ne modifient pas les caractéristiques des évolutions dues aux contraintes internes, de la même façon que ces dernières n'ont pas d'influence sur les caractéristiques des évolutions résultant des contraintes sociales*. Si des différences peuvent être observées, elles le sont plutôt au niveau quantitatif, comme nous l'avons vu plus haut.

L'ensemble de ces études est encore très préliminaire vis à vis de nombreux aspects, et ne peut offrir que des conclusions partielles et fragiles sur l'évolution des systèmes linguistiques. Un de nos projets dans un futur proche sera d'approfondir nos études, selon de multiples directions :

- en considérant d'autres types de paysage énergétique, en particulier des paysages présentant des reliefs plus riches ;
- en tentant de définir des indicateurs plus pertinents de l'évolution du système, permettant en particulier de caractériser la directionnalité des évolutions de celui-ci, élément qui a fait tout à fait défaut dans les expériences précédentes ;
- en testant des situations asymétriques pour les liens sociaux, et surtout en testant des situations de contact ;
- en tentant d'appliquer notre modèle à des systèmes réels comme des systèmes vocaliques ;
- en mettant en jeu la notion de contexte que nous n'avons pas exploitée dans nos expériences.

#### 6.2.4 Analyse et extrapolation des résultats des expériences

##### Idiolectes, langages communaux et agents

La première étape qu'il convient de franchir avant d'analyser plus finement les résultats précédents est d'opérer le lien entre les phénomènes réels que sont les idiolectes ou les langues avec les agents du système.

Le lien le plus immédiat est bien sûr celui qui fait correspondre à chaque agent un individu et son idiolecte. Les réseaux sociaux de nos expériences sont alors les réseaux sociaux qui unissent les individus dans la réalité.

Toutefois, comme le montre la figure 6.45, qui illustre l'évolution pendant 10,000 pas de temps d'un réseau composé de deux classes distinctes de 8 individus reliés par des liens positifs (les trajectoires bleue et blanche représentent les deux barycentres des deux classes, tandis que les trajectoires rouges et jaune représentent deux individus appartenant chacun à une des classes), les évolutions du barycentre d'un système linguistique sont très voisines de celles d'un des agents de ce même système. Elles apparaissent bien de plus faible amplitude et arborent un aspect plus tortueux, mais elles présentent néanmoins comme les mouvements des agents les caractéristiques d'un mouvement brownien. Dans le cas de liens négatifs, cette dernière propriété est également conservée, mais les agents et leur barycentre sont éloignés.

Sur la base des similarités précédentes que nous pourrions qualifier de structurelles, nous pouvons nous demander s'il n'est pas possible de considérer également dans certains cas un agent comme un système communal représentant un ensemble d'individus. Nous pensons que ceci est possible à partir du moment où les agents possèdent des idiolectes voisins, ce qui conduit à un barycentre représentatif des réalités linguistiques de la communauté en jeu.

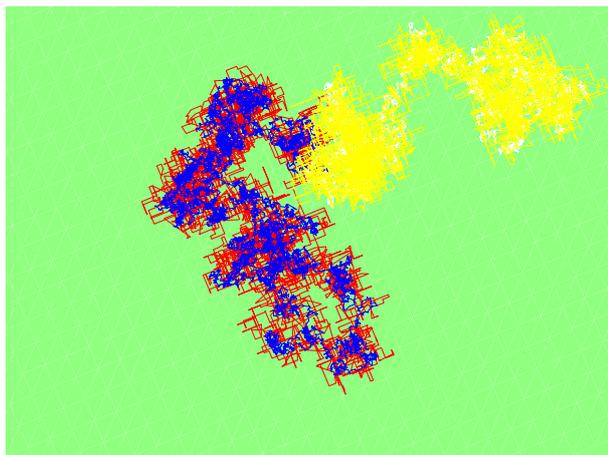


FIG. 6.45 – Evolution des barycentres des membres de deux classes différentes, et comparaison avec les trajectoires de deux agents de ces classes

Nous pouvons renforcer cet argument en insistant sur la dimension fractale des réseaux sociaux dans les communautés humaines. En effet, si de nombreux réseaux locaux sont formés par les interactions d'un nombre restreint d'individus, des macro-structures sociales plus importantes existent également et englobent les réseaux précédents, tout en étant englobées elle-mêmes dans des structures d'encore plus grande taille. Les plus globales et les plus connues sont vraisemblablement les classes sociales, sans compter les structures horizontales entre personnes d'une même classe d'âge, et les deux classes correspondant aux deux sexes.

Dès lors, il semble possible d'inscrire un agent dans cette macro-structure fractale, et de le faire correspondre à une certaine granularité de description pertinente pour l'expérience. Ceci signifie qu'un agent peut englober un nombre plus ou moins important d'individus selon le niveau de détail souhaité. Le plus simple pour définir l'agent en jeu est alors de lui assigner la valeur du barycentre des idiolectes qu'il regroupe, bien que nous ayons vu plus haut que cette démarche n'est pas toujours la plus pertinente.

Si un agent regroupe un ensemble d'individus, est-il possible de traduire au niveau de cet agent les principales caractéristiques des idiolectes qui le composent ? Nous pensons que jouer sur la valeur de la variance sociale peut traduire des caractéristiques du groupe social. Une valeur peu importante refléterait un groupe cherchant à s'identifier au reste de la communauté, et reproduisant fidèlement au niveau linguistique la norme communale. Au contraire, une valeur de la variance plus élevée traduirait une plus grande indépendance du groupe vis à vis de la norme du reste de la communauté.

En franchissant encore une étape dans l'abstraction de la notion d'agent, nous pouvons même envisager que des agents représentent des *langues* différentes. De la même façon que la linguistique historique postule une homogénéité de la population pour ses travaux, nous remplacerions alors l'ensemble des locuteurs de la langue par une seule entité. Encore une fois, nous pensons que ceci est rendu possible par la proximité des idiolectes des locuteurs d'une même langue. Une telle représentation permet alors d'étudier des situations de contacts linguistiques à un niveau macro-populationnel, même s'il nous paraît plus réaliste dans de telles situations de redescendre

à une échelle inférieure, afin de bien capturer la complexité des phénomènes en jeu.

En postulant les propositions précédentes, nous avons bien sûr en tête l'application de notre modèle aux structures sociétales de la préhistoire. Nous pouvons rappeler ici que le schéma sociétal de base pour cette période est un groupe de 25 individus environ.

L'intérêt de pouvoir considérer un agent comme un individu ou un groupe d'individus est que notre modèle peut dès lors traiter à la fois des évolutions internes à un groupe particulier, et celles d'un ensemble de groupes en interactions. Le cas préhistorique rend d'autant plus licite le double niveau d'interprétation du modèle que la faiblesse des contacts entre groupes crée une ligne de démarcation franche entre les contacts inter-groupes et les contacts intra-groupes (en termes d'intensité).

### Quel type de paysage énergétique pour les systèmes linguistiques ?

Lors de la conclusion du chapitre 2 et aussi au sein même de ce chapitre, nous avons explicitement posé la question de la topographie du paysage énergétique dans lequel évoluent les systèmes linguistiques. Une première proposition est un paysage plat sans optima, qui résulterait d'une trop grande faiblesse des contraintes internes sur les structures linguistiques pour créer un relief marqué. Une seconde proposition est un paysage au relief plus marqué, dans lequel les systèmes linguistiques peuvent évoluer sous l'action de diverses forces.

La première série d'expériences réalisée plus haut nous permet de nous pencher sur cette question avec une conception plus tangible des évolutions linguistiques d'un système selon les différents cas. La réflexion s'effectue sur un système en isolation, et nous observons différents scénarios évolutifs selon le choix de la variance naturelle :

- si la variance est nulle ou quasi-nulle, le système évolue en direction du minimum local dont le bassin d'attraction contient sa configuration initiale, puis se stabilise une fois ce minimum atteint ;
- si la variance n'est pas nulle mais reste inférieure à une certaine valeur critique, le système fluctue autour d'un attracteur au sein du bassin d'attraction de celui-ci, mais ne parvient jamais à franchir les barrières d'énergie autour de cet attracteur, ce qui lui permettrait de gagner une autre bassin ;
- pour des valeurs de la variance naturelle au-delà de la valeur limite précédente, mais qui demeurent néanmoins limitées, le système évolue autour des attracteurs, mais peut sauter de l'un à l'autre au gré d'événements contingents. Ces derniers lui permettent en effet, par une accumulation d'évolutions allant à l'encontre des contraintes internes, de franchir les barrières énergétiques séparant les attracteurs. De façon intuitive, plus une barrière est haute, plus le système a des difficultés à la franchir ; selon la hauteur, les sauts peuvent être très rares, voire impossibles ;
- enfin, pour des valeurs de la variance très importantes, le système évolue dans l'espace des paramètres sans réellement tenir compte de la topographie.

La première situation précédente semble en contradiction avec l'évolution des langues telle que nous l'observons aujourd'hui, puisque le modèle prédit un système final stable, alors que les systèmes linguistiques contemporains sont en perpétuelle évolution. Deux appréciations différentes du phénomène peuvent être mentionnées ici : une première possibilité est que les langues actuelles n'ont toujours pas fini leur course vers un minimum du paysage énergétique, et que

leurs évolutions servent à améliorer en permanence leur adéquation aux contraintes internes, c'est à dire à descendre les pentes du bassin d'attraction en direction de l'attracteur. Dès lors, nous devrions nous attendre à une stabilisation des langues dans un futur plus ou moins lointain. Il peut paraître étrange, au vu des évolutions observées en linguistique historique ou en sociolinguistique, que les langues aient pu évoluer depuis plusieurs milliers, voire dizaines de milliers d'années, sans attendre d'optimum local, mais cette question ne peut que rester sans réponse à cause de notre ignorance de la topologie de l'espace des états possibles.

Cette hypothèse nous paraît cependant difficile à accepter si l'on repense aux conditions du modèle d'émergence par polygenèse structurelle des stratégies linguistiques que nous avons développé au chapitre 4. En effet, selon ce modèle, les langues auraient dû se diversifier et se complexifier en adoptant d'emblée des structures et des items en bonne adéquation avec les contraintes naturelles. Dès lors, il est difficile d'expliquer comment nous aurions pu passer d'un stade initial en bonne adéquation avec les contraintes à un stade très déséquilibré dont nous nous écarterions progressivement aujourd'hui.

La seconde appréciation que nous pouvons donner du résultat précédent est qu'il reflète le paradoxe Saussurien que nous avons évoqué au chapitre 2. En effet, le poids des contraintes internes est ici si fort sur les structures linguistiques qu'il est impossible de s'éloigner d'un attracteur sans déséquilibrer trop fortement le système. Comme nous l'avons déjà dit, cette vision extrême de la structuration d'un système ne nous semble pas pertinente.

La seconde possibilité d'évolution d'un système, où la trajectoire évolutive ne parvient pas à quitter le bassin d'attraction d'un attracteur, correspondrait à une situation où l'on pourrait observer de la variabilité mais pas de diversité. Une fois encore, cette vision des choses, similaire aux réflexions de Saussure, ne semble pas plus en adéquation avec la réalité.

A l'inverse, la dernière situation, où une variance trop importante nie la topographie du paysage énergétique, conduit à une absence de toute stabilisation du système, et soulève la question de la dynamique des évolutions linguistiques réelles. Le problème est en effet ici de savoir si les évolutions des idiolectes ou des langues sont absolument libres, ce que suggère Mufwene (voir les derniers paragraphes du chapitre 2). Le constat d'évolution perpétuelle des langues semble bien nous mener en effet vers une suite de changements ininterrompue.

Cette vision des choses ne nous paraît pas correcte pour la raison suivante : si les langues évoluaient de façon totalement non contrainte, alors les typologistes ne devraient pas observer de schémas significatifs de fréquences de distributions des items de différents sous-systèmes linguistiques tels qu'ils existent dans les langues du monde, par exemple les systèmes phonologiques ou même les universaux implicationnels. L'existence de telles distributions ne nous semble pouvoir être expliquée que par l'existence de minima locaux du paysage énergétique.

Ceci nous mène à la dernière situation d'évolution, qui nous paraît la plus conforme à la réalité des changements linguistiques. Dans cette situation, le système passe des périodes de temps plus ou moins longues dans le voisinage d'un attracteur en fluctuant autour de lui, avant de franchir une barrière énergétique et de se rapprocher d'un autre attracteur.

Un attracteur sera d'autant plus stable que son bassin d'attraction sera pentu. Les fréquences des différents systèmes vocaliques peuvent ainsi s'expliquer par des attracteurs aux bassins d'attraction plus ou moins marqués [Vallée, 1994].

Une évolution d'un système selon cette dernière dynamique se réalise par **équilibres ponctuels**. De rapides périodes de transitions alternent avec de longues phases de relative stabilité

(avec des fluctuations autour d'un équilibre).

Ce schéma nous semble correspondre à la réalité des évolutions des langues. En effet, il nous semble que les sous-systèmes linguistiques restent stables pendant des périodes de temps assez prolongées, avant de voir des changements s'opérer en leur sein, et atteindre un nouvel état d'équilibre.

Si l'on considère l'exemple du *Great Vowel Shift* en Angleterre, cette étape montre une transition dans le système phonologique de la langue, tout comme par ailleurs le phénomène de nasalisation d'une partie des voyelles en français au cours du Moyen-Âge. Avant et après ces épisodes transitoires, les systèmes phonologiques des deux langues sont restés globalement les mêmes pendant une longue durée. Au niveau du lexique, des changements s'opéreront également pendant des phases de transition où par diffusion lexicale, les mots seront graduellement touchés par un changement phonétique. Les formes résultantes pourront alors rester stables pendant une longue période de temps. De la même façon enfin, des changements syntaxiques importants, comme l'importation de tons, le changement de l'ordre des mots etc., seront des événements rapides par rapport aux périodes de stabilité qui les encadreront.

Nous pouvons aussi souligner la congruence de ce scénario avec les propositions de Dixon sur l'évolution par équilibres ponctués des langues. Il est intéressant toutefois de noter ici que les arguments de ce dernier portent sur une échelle temporelle différente de celle que nous considérons dans notre modèle. Que ce soit lors des périodes de ponctuation ou de stabilité du modèle de Dixon, les évolutions linguistiques se font toujours, au niveau de la *micro-histoire*, par équilibres ponctués, mais peuvent apparaissent graduelles à une échelle d'observation supérieure comme la *méso-histoire* (lors des périodes de stabilité dans le modèle de Dixon). En outre, les transitions abruptes que nous postulons pour notre modèle ne sont jamais qu'une suite d'évolutions continues et graduelles à une échelle de temps plus petite. Tout est affaire de compression du temps.

Il nous reste à résoudre l'apparent paradoxe d'un sentiment de perpétuelle évolution des systèmes linguistiques, alors que nos propositions vont plutôt dans le sens d'équilibres ponctués. Nous devrions donc logiquement observer dans ce dernier cas des périodes de stabilité sans aucun changement, séparées par des périodes beaucoup plus courtes de transformations structurelles.

Nous pensons que la résolution de ce paradoxe provient du grand nombre de sous-systèmes et de structures qui s'entrelacent dans la macro-structure qu'est le langage. L'impression de changement continu est causé selon nous par le fait que parmi le grand nombre d'items et de structures qui peuvent évoluer selon le processus précédent, il existe toujours un petit nombre d'entre eux qui varie à une date donnée, même si ces évolutions et transitions se produisent sous l'action de gâchettes sociales contingentes et externes au système (nous reprenons ici partiellement les propositions de Maturana). La relative indépendance des sous-systèmes linguistiques et des structures permet à des changements de se produire dans l'un d'entre eux sans modifier les items des autres. Le simple grand nombre de structures qui peuvent évoluer et qui évoluent effectuent aléatoirement conduit à une perpétuelle évolution du langage.

### **Système isolé ou système en contact**

La réflexion précédente repose entièrement sur la possibilité de pouvoir isoler un système linguistique d'autres systèmes.

Cette possibilité est difficile à envisager dans la réalité, puisque les idiolectes et les langues sont tous plus ou moins fortement en interaction, du fait des contacts des individus qui les sous-tendent. Le seul système linguistique réellement isolé correspond au macro-système qui recouvre l'ensemble des idiolectes de *tous* les individus. Cependant, les divergences importantes

des langues du monde entre elles font que leur barycentre, c'est à dire le macro-système précédent, ne possède pas une grande pertinence vis à vis de la réalité linguistique.

Néanmoins, nous pensons que le scénario d'évolution par équilibres ponctués proposé dans les paragraphes précédents n'est pas remis en cause par la présence de contraintes sociales. En effet, comme nous l'avons vu, il semble que les processus d'évolutions induits par les contraintes internes et ceux engendrés par les contraintes sociales sont opérationnellement distincts, au sens où nous avons défini cette expression plus haut. Dès lors, le scénario d'évolution par équilibres ponctués, qui résulte de l'existence de contraintes internes au système, n'est pas modifié qualitativement (mais seulement quantitativement) par la présence de contraintes externes. Vu sous un autre angle, ce ne sont pas ces dernières qui sont responsables de la dynamique d'évolution par équilibres ponctués. Ceci est confirmé expérimentalement par l'absence d'équilibres ponctués dans la deuxième série d'expériences, et par leur présence dans les première et troisième séries.

### Probabilité d'émergence et de stabilité de Greenberg

Comme nous l'avons souligné plus haut, notre modèle ne prend pas en compte explicitement la facilité d'émergence et la stabilité des items linguistiques comme l'a fait Greenberg lors de ses études typologiques. Toutefois, il est possible de déterminer assez simplement les configurations topographiques du paysage énergétique qui correspondent aux quatre cas cités par Greenberg [Greenberg, 1978] (p. 76) (voir chapitre 2).

Un élément très stable correspond à un puits assez profond et aux parois assez verticales, ce qui rend une sortie de ce puits difficile et donc rare. Au contraire, une forme instable correspondra à un puits de faible profondeur et aux pentes douces. En ce qui concerne la probabilité d'émergence, une forme qui émergera facilement dans les idiolectes des locuteurs (et donc dans une langue) correspondra à un puits facile d'accès, quand une forme apparaissant rarement correspondra elle à un puits difficile d'accès. La figure 6.46 présente des exemples simples de ces différentes situations.

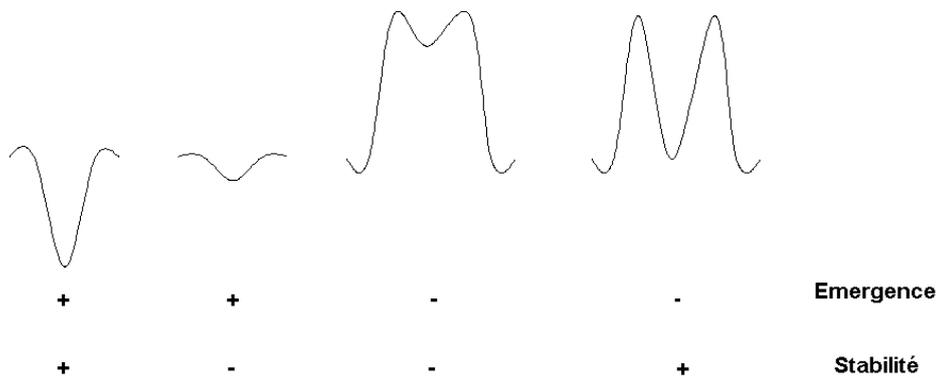


FIG. 6.46 – Différentes topologies de l'espace énergétique aptes à rendre compte des quatre schémas de stabilité et d'émergence de Greenberg [Greenberg, 1987] (p. 76)

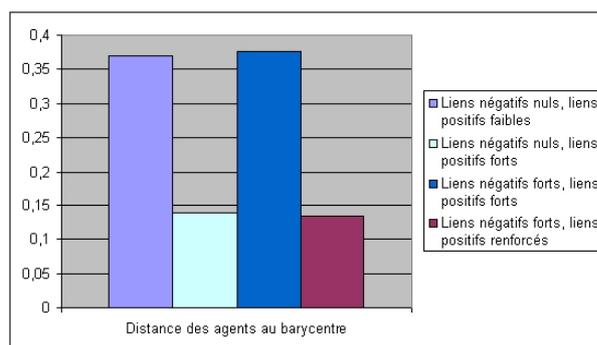


FIG. 6.47 – Distance moyenne au barycentre au cours des 4 phases du scénario d'évolution des structures sociétales

### Retour sur quelques considérations à propos de la préhistoire des langues

Pour finir ce chapitre, nous souhaitons présenter une expérience rudimentaire résumant les précédentes et permettant de revenir sur notre scénario global d'évolution des langues au cours de la préhistoire. Cette expérience se base sur un paysage énergétique plat, mais nous avons plus haut que ceci conduit aux mêmes résultats qu'un paysage au relief marqué. La distance des agents au barycentre constitue en particulier un indicateur valide de la diversité linguistique du macro-système.

Dans cette expérience, 10 systèmes linguistiques sont reliés par un réseau de liens plus ou moins dense en terme de connexions positives ou négatives. Plus concrètement, un paramètre de densité des connexions positives correspond au nombre de liens de valeur 0.05 qui sont créés aléatoirement pour relier un agent aux autres membres de la communauté. Ces liens peuvent s'ajouter les uns aux autres, et donc plus la densité importante, et plus l'agent est relié fortement par des liens attractifs aux autres agents. Un paramètre de densité des connexions négatives permet de manière similaire de créer aléatoirement un certain nombre de connexions négatives.

Nous envisageons successivement quatre situations : dans la première, la densité de liens négatifs est nulle, et celle de liens positifs très faible (5). Dans le second cas, la densité de liens négatifs est toujours nulle, mais celle de liens positifs est plus importante (625). Dans la troisième situation, la densité de liens négatifs croît jusqu'à égaler la densité de liens positifs (625). Enfin, une dernière étape consiste à augmenter une dernière fois le nombre la densité de liens positifs (3125), en gardant la seconde densité inchangée.

Les figures 6.47, 6.48, 6.49 représentent comme à l'habitude la distance des agents au barycentre, l'évolution de la distance moyenne parcourue par les agents et celle du barycentre du système.

Nous envisageons les systèmes linguistiques manipulés dans la simulation comme les systèmes communaux des différents groupes humains en interaction au cours de la préhistoire. Nous observons que nos prédictions du chapitre 4 quant aux évolutions de la diversité linguistique des groupes humains au cours de la préhistoire sont bien vérifiées par la simulation : une réduction de la diversité linguistique s'opère avec le premier accroissement des liens positifs, mais cette réduction disparaît lors de l'augmentation de la densité de liens négatifs liée à l'accroissement des densités de population à la fin du Paléolithique Supérieur et au Néolithique. Enfin, la der-

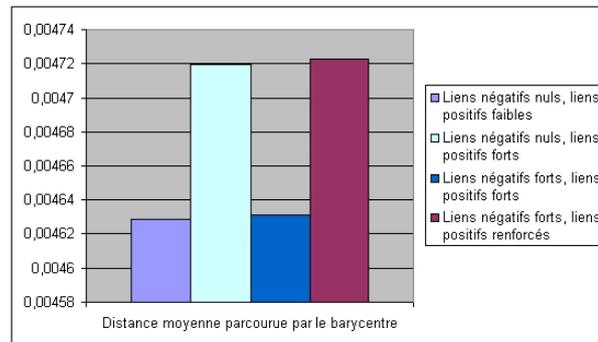


FIG. 6.48 – Distance moyenne parcourue par les agents au cours des 4 phases du scénario d'évolution des structures sociétales

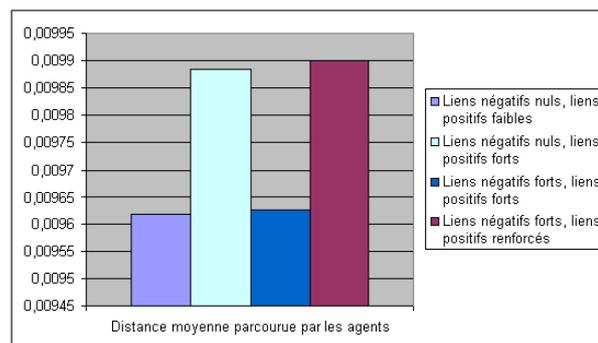


FIG. 6.49 – Distance moyenne parcourue par le barycentre du macro-système au cours des 4 phases du scénario d'évolution des structures sociétales

nière étape montre comment une nouvelle augmentation de la densité de liens positifs entraîne une nouvelle diminution de la diversité linguistique, ce qui semble correspondre à la situation contemporaine où de nombreuses langues disparaissent.

L'intérêt de notre modèle est qu'il nous renseigne également sur les modifications du rythme d'évolution des agents et du macro-système dans son ensemble.

Nous constatons une évolution de ces rythmes d'évolution à l'inverse de celle de la diversité linguistique, mais de façon très peu marquée. Lors de la première phase de notre scénario, une forte diversité correspond à un faible rythme d'évolution des agents et du macro-système. La décroissance de la diversité au cours de la seconde étape s'accompagne d'un très léger accroissement des rythmes d'évolution, qui diminuent à nouveau légèrement lors de la transition Néolithique, pour croître une dernière fois de quelques pourcents lors de la quatrième phase du scénario. Ces faibles variations sont en accord avec les résultats établis plus haut dans la troisième série d'expérience, et semblent indiquer que la vitesse d'évolution des langues au cours de la préhistoire n'a été que peu marqué par les variations des liens sociaux, que ces langues aient été plus ou moins voisines typologiquement (diversité plus ou moins importante). Néanmoins, il convient de demeurer prudent sur ce résultat, qui pourrait être un biais du modèle lié au calcul de la déviation sociale optimale. Des expériences jouant sur le procédé même de calcul des interactions sociales sont nécessaires avant de juger si la grande stabilité des vitesses d'évolution en fonction de la force des liens est un artéfact ou une réalité.

Notons ici que ces évolutions ne prennent pas en compte les pressions d'exploration de l'espace des possibles qui ont pu exister après une modification de celui-ci suite à l'apparition de nouvelles capacités cognitives et/ou physiologiques. Il faut donc voir les courbes précédentes plutôt comme des modulations *sociales* à apporter à la dynamique générale *interne* d'exploration de l'espace des possibles.

Les résultats de cette expérience sont bien sûr très grossiers, et ne prennent pas en compte de façon fine plusieurs phénomènes qui s'entrecroisent lors de l'évolution des structures sociales, à savoir l'accroissement de la taille des groupes humains, la possible modification des structures sociales en leur sein, lors du Néolithique par exemple, ou encore l'étalement géographique des liens. Ces phénomènes, qui prennent place selon nous surtout lors de la fin du Paléolithique Supérieur et du Néolithique, viennent moduler les résultats précédents et rendent la situation plus délicate à concevoir.

## 6.3 Conclusions

Nous nous sommes efforcés dans ce chapitre de présenter et d'étudier un modèle général des évolutions et des changements linguistiques. Ce modèle se démarque en particulier des modèles plus traditionnels par la prise en compte spécifique et l'étude de l'impact des liens sociaux entre les locuteurs ou les communautés linguistiques.

Différentes expériences nous ont permis de mieux estimer l'impact des contraintes internes et externes sur les dynamiques d'évolution d'un système linguistique. En particulier, nous avançons l'hypothèse d'une évolution des systèmes linguistiques selon une dynamique par **équilibres ponctués**. En mettant en évidence une **disjonction opérationnelle** des influences des contraintes internes et externes du système, nous émettons l'idée que cette dynamique d'évolution ne dépend que des contraintes internes au système.

Plus généralement, les résultats des expériences laissent entrevoir la sensibilité de la diversité et de la vitesse d'évolution des systèmes linguistiques aux structures sociales. Ceci semble

rejoindre les critiques envers la glottochronologie, qui postule pour toutes les langues un taux de changement constant pour les mots du vocabulaire de base. Néanmoins, il reste à déterminer si l'inertie de ce vocabulaire peut-être telle que les différences sociolinguistiques n'aient pas d'emprise sur lui.

Nous avons enfin proposé une validation informatique (bien que rudimentaire) des hypothèses avancées au chapitre 4 sur l'évolution de la diversité linguistique au cours de la préhistoire.

Nous espérons par ce dernier chapitre compléter les deux précédents, et parvenir à clore une approche la plus riche possible de l'émergence et de l'évolution des systèmes linguistiques au cours du passé. Il va néanmoins de soi que de nombreux développements sont à apporter au modèle : tests de structures sociales plus riches ou de paysages énergétiques plus complexes, recherche d'une meilleure définition mathématique des contraintes sociales. . .