Annexe A : Approche combinatoire et représentation matricielle du modèle d'évolution linguistique

Dans cette annexe, nous souhaitons développer rapidement le traitement mathématique du modèle développé au chapitre 6 selon une approche non holistique. L'approche combinatoire paraît en effet plus facile à appliquer à des cas réels, puisqu'au lieu de devoir définir une fonction d'énergie sous forme analytique, il suffit de fournir au système les compatibilités intrinsèques et relationnelles des différents items en jeu. Ce sont ces données qui permettront de construire la fonction d'énergie qui sera manipulée par le modèle.

Le formalisme utilisé est un formalisme matriciel, qui se prête particulièrement bien à des implémentations informatiques.

3.1 Représentation matricielle des contraintes

Le pool d'items va ici être défini par une matrice $\mathbf{C}$ dont les valeurs vont servir à représenter les compatibilités entre les différents items.

En supposant un ensemble de $\mathbf{n}$ items, la ligne $\mathbf{i}$ de la matrice représentera les compatibilités relationnelles de l'item $\mathbf{i}$ avec les autres items. Nous adoptons la définition suivante : l'élément $(\mathbf{i}, \mathbf{j})$ de la matrice représente la compatibilité relationnelle (dénommée par **CR** ci-dessous) de l'item $\mathbf{i}$ vis à vis des contraintes dans un système linguistique contenant l'item $\mathbf{j}$ (au sens défini au chapitre 6). Les valeurs adoptées pour la **CR** varient sur l'ensemble des réels entre $-\infty$ et $+\infty$:

- une valeur positive pour l'élément $(\mathbf{i}, \mathbf{j})$ signifie que l'élément $\mathbf{i}$ est compatible avec le système contenant l'item $\mathbf{j}$: les contraintes sont satisfaites d'une façon globale. Ce caractère global, et le résumé de l'ensemble des contraintes dans quelques valeurs numériques, signifie par exemple au niveau phonétique que la perception est facilité et la production rendue plus difficile, mais que *globalement* la présence de l'item $\mathbf{i}$ rend l'utilisation du système linguistique plus aisée ;
- une valeur négative signifie au contraire que l'élément $\mathbf{i}$ est incompatible en présence de l'item $\mathbf{j}$; il va à l'encontre du jeu de contraintes ;
- une valeur nulle signifie que la compatibilité de l'item $\mathbf{i}$ vis à vis des contraintes n'est pas influencée par la présence dans le système de l'item $\mathbf{j}$.

Plus la valeur de la **CR** est grande en valeur absolue, plus la compatibilité ou l'incompatibilité est importante, ce qui avec notre définition des contraintes signifie que l'utilisation du langage est plus fortement facilitée ou au contraire entravée.

La définition précédente s'applique pour tous les éléments $(\mathbf{i}, \mathbf{j})$ de la matrice où $\mathbf{i}$ est différent de $\mathbf{j}$. Les éléments $(\mathbf{i}, \mathbf{i})$ de la matrice sont eux choisis pour mesurer la compatibilité intrinsèque de l'item $\mathbf{i}$, et varient eux aussi entre $-\infty$ et $+\infty$.

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

$$\mathbf{C}[\mathbf{i}, \mathbf{j}] = \mathbf{C}_{i,j} = \text{compatibilité de l'item } \mathbf{i} \text{ vis à vis des contraintes en présence de l'item } \mathbf{j}$$

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

Comme pour l'approche holistique dans le chapitre 6, les contextes vont être représentés à l'aide de vecteurs de dimension $\mathbf{n}$, chaque ligne du vecteur correspondant à la fréquence d'un item $\mathbf{i}$ dans ce contexte.

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

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

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

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

3.3 Définition de la compatibilité systémique d'un item et de la cohérence systémique d'un système

Un des intérêts d'associer compatibilités intrinsèques et relationnelles des items linguistiques dans une seule matrice est de permettre de calculer la compatibilité systémique d'un item dans un contexte particulier ou un ensemble de contextes. A partir du vecteur des fréquences des items dans un contexte, nous définissons la compatibilité systémique $\mathbf{c_s}$ d'un item linguistique i dans le contexte $\mathbf{V}$ selon la formule :

$$\mathbf{c_s}(\mathbf{i}, \mathbf{V}) = \mathbf{f}_i \times \left(\mathbf{C}[\mathbf{i}, \mathbf{i}] + \sum_{j=1, j \neq i}^{\mathbf{N}} \mathbf{f}_j \times \mathbf{C}[\mathbf{i}, \mathbf{j}] \right)$$

La compatibilité systémique d'un item dans un contexte prend en compte d'une part sa fréquence d'occurrence dans le contexte, et d'autre part les fréquences d'utilisation des autres items qui s'expriment dans le contexte en les reliant aux compatibilités relationnelles. La composition de la formule de calcul se justifie de la façon suivante :

- plus la fréquence de l'item considéré dans le contexte est importante, plus l'effet de sa compatibilité systémique vis à vis des contraintes sera marqué. Un item de fréquence nulle a bien une compatibilité systémique nulle puisque n'apparaissant pas, il ne vient jouer aucun rôle vis à vis des contraintes. A l'opposé, plus la fréquence d'occurrence d'un item est importante, plus il joue un rôle important dans l'utilisation du langage, et plus sa compatibilité systémique peut prendre des valeurs importantes, négatives ou positives ;
- en ce qui concerne les fréquences d'occurrence des autres items, le raisonnement est symétrique : si la fréquence d'un item $\mathbf{j}$ est nulle, la compatibilité relationnelle de l'item $\mathbf{i}$ avec lui sera nulle, c'est à dire n'aura aucune influence sur la compatibilité de $\mathbf{i}$ vis à vis des contraintes. A l'opposé, plus l'item $\mathbf{j}$ aura une fréquence importante, plus il se manifestera souvent dans l'utilisation du système linguistique, et plus il sera nécessaire de prendre en compte sa relation avec $\mathbf{i}$ vis à vis des contraintes.

Remarquons que l'on peut décomposer la compatibilité systémique d'un item comme le produit de sa fréquence par sa **compatibilité systémique absolue**, c'est à dire la valeur de sa compatibilité systémique lorsque sa fréquence est égale à $\mathbf{1}$.

Le calcul de la compatibilité systémique d'un item est indépendante de la présence effective de celui-ci dans le contexte. Il est donc possible de mesurer cette valeur pour une fréquence quelconque et un système donné, ce qui permet d'envisager comment une modification fréquentielle augmente ou diminue la compatibilité systémique d'un item.

Il nous faut maintenant définir la cohérence systémique d'un contexte à partir des compatibilités systémiques des items qui le composent. Cette définition peut se faire simplement de la façon suivante :

$$\mathbf{c_s}(\mathbf{V}) = \frac{1}{\mathbf{n}} \sum_{j=1}^{\mathbf{n}} \mathbf{c_s}(\mathbf{j}, \mathbf{V})$$

La cohérence du système est la moyenne des compatibilités systémiques des items qui le composent (avec un jeu de fréquences donné. Si une fréquence est nulle, les formules montrent bien que l'item ne joue aucun rôle dans le calcul de la cohérence).

La cohérence systémique d'un système linguistique nous donne accès à la fonction d'énergie du système, puisque nous avons défini dans le chapitre 2 l'énergie d'un système linguistique comme son adéquation au jeu de contraintes naturelles et distribuées. Plus la cohérence systémique d'un système sera importante, et plus l'énergie sera faible, et inversement. L'énergie est donc simplement l'inverse de la cohérence systémique.

Il est possible d'étendre facilement les calculs de compatibilité systémique et de cohérence d'un système $\mathbf{S}$ dans le cas de plusieurs contextes indépendants $(\mathbf{V_1}, \ldots, \mathbf{V_q})$ ($\mathbf{S}$ est constitué de l'ensemble de ces contextes) :

$$\mathbf{c_s}(\mathbf{i}, (\mathbf{V_1}, \ldots, \mathbf{V_q})) = \frac{1}{\mathbf{q}} \sum_{k=1}^{\mathbf{q}} \mathbf{c_s}(\mathbf{i}, \mathbf{V_k})$$

$$\mathbf{c_s}(\mathbf{S}) = \sum_{k=1}^{\mathbf{q}} \mathbf{c_s}(\mathbf{V_k})$$

3.4 Dérivée de la fonction d'énergie

Pour pouvoir appliquer les mécanismes d'évolution spécifiés au chapitre 6, il est nécessaire de pouvoir calculer les dérivées partielles de la fonction d'énergie du système selon les différents paramètres du système.

Cette opération est en fait simple à réaliser au niveau mathématique, comme le résument les formules suivantes :

$$\begin{aligned}
\mathbf{c_s}(\mathbf{V}) &= \frac{1}{\mathbf{n}} \sum_{i=1}^{\mathbf{n}} \mathbf{c_s}(i, \mathbf{V}) \\
&= \frac{1}{\mathbf{n}} \sum_{i=1}^{\mathbf{n}} \left(f_i \times \left(\mathbf{C}[i,i] + \sum_{j=1, j\neq i}^{\mathbf{n}} f_j \times \mathbf{C}[i,j] \right) \right) \\
&= \frac{1}{\mathbf{n}} \Bigg(f_{i_0} \times \mathbf{C}[i_0, i_0] + \sum_{i=1, i\neq i_0}^{\mathbf{n}} f_i \times \mathbf{C}[i,i] + \sum_{j=1, j\neq i_0}^{\mathbf{n}} f_{i_0} \times f_j \times \mathbf{C}[i_0, j] \\
&\quad + \sum_{i=1, i\neq i_0}^{\mathbf{n}} f_i \times f_{i_0} \times \mathbf{C}[i, i_0] + \sum_{i=1, i\neq i_0}^{\mathbf{n}} \sum_{j=1, j\neq i, j\neq i_0}^{\mathbf{n}} f_i \times f_j \times \mathbf{C}[i,j] \Bigg)
\end{aligned}$$

$$\Longrightarrow \quad \frac{\partial \mathbf{c_s}(\mathbf{V})}{\partial f_i} = \frac{1}{\mathbf{n}} \left(\mathbf{C}[i,i] + \sum_{j=1, j\neq i}^{\mathbf{n}} f_j \times \mathbf{C}[i,j] + \sum_{j=1, j\neq i}^{\mathbf{n}} f_j \times \mathbf{C}[j,i] \right)$$

Les calculs précédents concerne un unique contexte, mais il est facile de les étendre à un ensemble de contextes. L'algorithme de *steepest descent* portera dès lors sur l'ensemble des fréquences dans les différents contextes. Les formules suivantes prennent en compte l'ensemble des contextes :

$$
\begin{aligned}
\mathbf{c_s}((\mathbf{V}_1, \ldots, \mathbf{V_q})) &= \frac{1}{\mathbf{n} \times \mathbf{q}} \sum_{k=1}^{\mathbf{q}} \sum_{i=1}^{\mathbf{n}} \mathbf{c_s}(i, \mathbf{V}) \\
&= \frac{1}{\mathbf{n} \times \mathbf{q}} \sum_{k=1}^{\mathbf{q}} \sum_{i=1}^{\mathbf{n}} \left(f_i \times \left(\mathbf{C}[i,i] + \sum_{j=1, j \neq i}^{\mathbf{n}} f_j \times \mathbf{C}[i,j] \right) \right)
\end{aligned}
$$

$$
\frac{\partial \mathbf{c_s}((\mathbf{V}_1, \ldots, \mathbf{V_q}))}{\partial f_i(\mathbf{V}_k)} = \frac{1}{\mathbf{n} \times \mathbf{q}} \left(\mathbf{C}[i,i] + \sum_{j=1, j \neq i}^{\mathbf{n}} f_j(\mathbf{V}_k) \times (\mathbf{C}[i,j] + \mathbf{C}[j,i]) \right)
$$