

UNE MESURE DE L'INTEGRATION D'UN ELEMENT A UN SYSTEME

Quelques mots pour rappeler sur quels phénomènes linguistiques nous travaillons à Lausanne et quels sont les résultats acquis¹. Les structures d'une langue ne reposent pas seulement sur des règles absolues mais aussi sur des régularités statistiques. C'est une évidence si l'on admet qu'une langue est toujours en cours de transformation. C'est aussi une nécessité fonctionnelle puisque la communication linguistique se réalise dans des situations qui ne sont pas totalement prédéterminées. Nous avons consacré nos recherches à l'étude de cette dimension de "rigueur" de la structuration linguistique. Soulignons, pour éviter tout malentendu, que la "variabilité linguistique" qui nous intéresse est donc celle qui s'observe dans une communauté linguistique homogène et non pas celle qu'on peut corrélérer avec des variables sociales externes en sociolinguistique. Dans une première étape nous avons mis en évidence l'existence du phénomène aux deux niveaux d'observation de la réalité linguistique: celui du groupe et celui de l'individu.

Dans l'étape actuelle nous essayons de dégager des propriétés des phénomènes linguistiques qui permettraient de rendre compte de la rigueur de leur structuration. Propriétés relatives aux conditions d'occurrence de ces phénomènes et propriétés structurelles. Nous regroupons les premières sous le terme (à prendre comme une désignation encore vague et assez arbitraire) de "fréquence". Par hypothèse, plus un phénomène est

"fréquent" plus il est rigoureusement structuré². Quant aux propriétés structurelles, nous tentons de les contrôler à travers le concept d'"intégration" qui les résume. Par hypothèse, plus un phénomène est intégré plus il est rigoureusement structuré. En combinant "fréquence" et intégration et en s'en tenant aux deux pôles de la structuration linguistique (supposée continue) on prévoira qu'un phénomène "fréquent" et/ou bien intégré relève des structures les plus rigoureuses alors qu'un phénomène non "fréquent" et mal intégré relève des structures les moins rigoureuses.

Nous ne discuterons ici que de la notion d'intégration. Elle dépend à l'évidence de celle de système, qu'elle nuance. Dès son utilisation en phonologie diachronique³ - dans l'hypothèse, largement confirmée, selon laquelle l'intégration d'un phonème à un système le met relativement à l'abri d'une évolution spécifique - le terme recouvre non l'idée d'une dichotomie (intégré/non intégré) mais celle d'une échelle: on parle de degrés d'intégration⁴. Il doit donc être possible, au moins dans certains cas et en principe, de hiérarchiser les éléments d'un système selon leur intégration plus ou moins forte à ce système. C'est ce que nous entendons par "mesurer" l'intégration. Des propositions ont déjà été faites dans cette voie. En phonologie, on a suggéré de faire reposer cette mesure sur le nombre de corrélations (oppositions bilatérales) auxquelles participe un phonème ou sur le nombre de proportions (oppositions bilatérales et multilatérales) dans lesquelles il entre. Mortéza Mahmoudian, qui généralise la notion d'intégration et envisage de

l'appliquer à n'importe quel système défini comme un ensemble d'entités et un ensemble de propriétés (donc représentable par un tableau à double entrée), a proposé deux mesures nouvelles⁵. Une entité serait d'autant plus intégrée qu'elle présente davantage de propriétés caractéristiques du système: c'est la mesure par la "complexité". Une entité serait d'autant plus intégrée qu'il y a plus d'entités dans chacune des classes dont le produit logique détermine son identité: c'est la mesure par la "solidarité". Si un système phonologique comporte 3 phonèmes caractérisés comme /bilabiales/, 4 comme /sourdes/ et 8 comme /orales/ et si /p/ y est défini comme /bilabial/, /sourd/, /oral/, son degré d'intégration sera de 3 du point de vue de la complexité et de 15 (3+4+8) du point de vue de la solidarité.

L'application des différentes mesures proposées montre qu'elles sont sensibles au choix de l'ensemble des propriétés qui caractérisent le système, ce qui est normal, mais aussi qu'elles sont, la solidarité mise à part, peu discriminantes. Partons de l'ensemble de phonèmes /p/, /b/, /m/, /t/, /d/, /n/, /k/ et /N/ (qui note la nasale vélaire) pour lequel on proposera deux analyses. Dans l'une (système A), seul /k/ est /oral/, ce trait étant impliqué par /sourd/ ou /sonore/ pour les autres occlusives non nasales; dans l'autre (système B), toutes ces occlusives non nasales sont caractérisées comme /orales/. Voici quelles sont les valeurs des degrés d'intégration de ces phonèmes selon les quatre mesures proposées et pour les deux systèmes:

Systèmes:	Corrélations		Proportions		Complexité		Solidarité	
	A	B	A	B	A	B	A	B
/p/	2	2	4	4	2	3	3	7
/b/	2	2	4	4	2	3	3	7
/m/	2	1	4	3	2	2	4	4
/t/	2	2	4	4	2	3	3	7
/d/	2	2	4	4	2	3	3	7
/n/	2	1	4	3	2	2	4	4
/k/	1	1	3	3	2	2	1	5
/N/	1	1	3	3	2	2	3	3

Il ressort de ce tableau que, dans le système B, trois mesures sur quatre donnent une même hiérarchie, qui ne comporte que deux degrés. Dans le système A, deux mesures aboutissent à une même hiérarchie à deux degrés tandis qu'une troisième ne détermine aucune hiérarchie. Finalement, c'est la mesure fondée sur la solidarité qui paraît la plus efficace. Dans l'ensemble, on ne peut pas dire que ces résultats soient particulièrement encourageants...

Pour tenter d'améliorer cette situation nous nous proposons de travailler d'abord sur l'ensemble des propriétés qui définissent le système, dans l'idée d'associer à chacune une mesure de ce que l'on pourrait appeler son "pouvoir structurant". Mahmoudian a fait deux propositions* qui vont nous servir de point de départ. Il distingue ce que nous nommerons la "centralité" des propriétés (qu'il appelle co-applicabilité) de leur "généralité" (qu'il appelle fréquence paradigmatique).

Une propriété est d'autant plus centrale que la somme de ses intersections avec chacune des autres propriétés comporte plus d'entités. Une propriété est d'autant plus générale qu'elle est caractéristique de plus d'entités. Ces deux aspects sont indépendants comme en témoigne la figure donnée en annexe I, où les trois rectangles correspondent à trois propriétés et les croix à des entités qui présentent les propriétés correspondant aux rectangles dans lesquels elles sont incluses.

Pour établir une mesure de l'intégration nous commençons par mesurer la centralité et la généralité de toutes les propriétés du système. Nous associons à chaque entité du système une valeur égale à la somme des valeurs de centralité ou de généralité des propriétés qui la caractérisent. Ces principes appliqués aux systèmes phonologiques qui nous ont servi d'exemple donnent les résultats suivants:

Systèmes:	Centralité		Généralité	
	A	B	A	B
/p/	5	18	5	10
/b/	5	18	5	10
/m/	6	8	6	6
/t/	5	18	5	10
/d/	5	18	5	10
/n/	6	8	6	6
/k/	3	11	3	7
/N/	5	5	5	5

Centralité et généralité aboutissent donc, dans chaque

système, à la même hiérarchie, celle que nous avait déjà fournie la prise en compte de la solidarité. Mais ce résultat est lié au statut particulier des tableaux phonologiques, tableaux de définitions d'entités nécessairement distinctes. Si l'on se propose de traiter des tableaux plus généraux, par exemple un ensemble d'éléments lexicaux caractérisés par un ensemble de propriétés syntaxiques, la mesure de l'intégration à partir de la généralité risque d'être fortement influencée par les entités définies par une seule et même propriété lorsque celles-ci sont nombreuses (cf. annexe II). C'est pourquoi nous calculerons l'intégration d'une entité en faisant la somme des valeurs de centralité des propriétés qui la caractérise.

Pour apprécier l'intérêt d'une telle mesure nous pouvons avoir recours d'abord à notre intuition. Appliquée à la fameuse analyse en sèmes des noms de sièges⁷ on aboutit à la hiérarchie suivante, du plus intégré au moins intégré (la valeur mesurant l'intégration figure entre parenthèses): chaise (69), tabouret (66), fauteuil (61), canapé (44) et pouf (44). Il nous semble que ce classement n'est pas quelconque et qu'on pourrait le mettre en relation avec la notion d'objet typique: comme siège la chaise est plus typique que le pouf de même que l'hirondelle est un oiseau plus typique que le pingouin. Mais ce n'est que lorsque nous disposerons de propositions précises pour la mesure de la "fréquence", au sens évoqué ci-dessus, qu'en combinant "fréquence" et intégration nous pourrions réellement valider l'une et l'autre: permettront-elles de retrouver la hiérarchie dans la rigueur de structuration que nous

avons pu observer aux deux niveaux, social et individuel, de la réalité linguistique? Autant dire qu'il reste du pain sur la planche... et pas mal de problèmes. Les plus importants sont évidemment liés à l'ensemble des propriétés qui définissent un système. Quelles propriétés faut-il retenir et ont-elles toutes la même importance? Autant de question qui se posent déjà pour des systèmes aussi simples que les systèmes phonologiques ce qui laisse présager d'énormes - d'insurmontables? - difficultés.

Université de Lausanne

Remi Jolivet

Notes

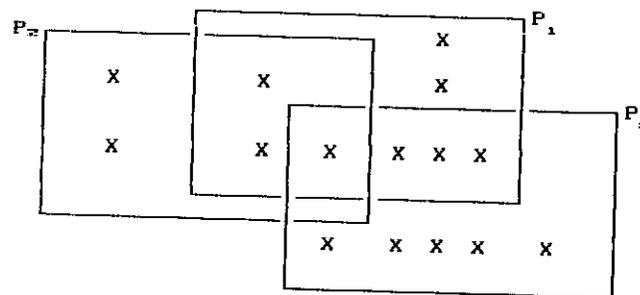
1. Pour une présentation d'ensemble cf. La linguistique, 16, 1, 1980, p. 5-117; résumé des recherches empiriques dans: R. Jolivet, Aspect statistique de la structuration linguistique, in: Méthodes quantitatives et informatiques dans l'étude des textes, Genève, Paris, Slatkine, Champion, 1986, p. 507-518. La question présentée dans cette communication est discutée de façon plus détaillée dans R. Jolivet, Mesurer l'intégration?, La linguistique, 22, 2, 1986, p. 3-19.
2. Contre-exemple évident en sémantique: en règle générale, plus un terme est fréquent, au sens statistique, plus son sens est flou, variable. Le contre-exemple n'est peut-être qu'apparent mais il a le mérite de souligner que toutes ces notions ne sont pas simples.
3. Cf. André Martinet, Economie des changements phonétiques, Berne, A. Francke, 1955, 396 p.
4. Ibid., 3.28
5. Mortéza Mahmoudian, Structure linguistique: problèmes de la constance et des variations, La linguistique, 16, 1, 1980, cf. p. 20-23
6. Ibid.
7. Cf. Bernard Pottier, Recherches sur l'analyse sémantique en linguistique et en traduction mécanique, Publications de la Faculté des Lettres et Sciences humaines de Nancy, Nancy, 1963, cité d'après Christian Baylon, Paul Fabre, La sémantique, Paris, Nathan, 1978, p. 75

Notes

1. Pour une présentation d'ensemble cf. La linguistique, 16, 1, 1980, p. 5-117; résumé des recherches empiriques dans: R. Jolivet, Aspect statistique de la structuration linguistique, in: Méthodes quantitatives et informatiques dans l'étude des textes, Genève, Paris, Slatkine, Champion, 1986, p. 507-518. La question présentée dans cette communication est discutée de façon plus détaillée dans R. Jolivet, Mesurer l'intégration?, La linguistique, 22, 2, 1986, p. 3-19.
2. Contre-exemple évident en sémantique: en règle générale, plus un terme est fréquent, au sens statistique, plus son sens est flou, variable. Le contre-exemple n'est peut-être qu'apparent mais il a le mérite de souligner que toutes ces notions ne sont pas simples.
3. Cf. André Martinet, Economie des changements phonétiques, Berne, A. Francke, 1955, 396 p.
4. Ibid., 3.28
5. Mortéza Mahmoudian, Structure linguistique: problèmes de la constance et des variations, La linguistique, 16, 1, 1980, cf. p. 20-23
6. Ibid.
7. Cf. Bernard Pottier, Recherches sur l'analyse sémantique en linguistique et en traduction mécanique, Publications de la Faculté des Lettres et Sciences humaines de Nancy, Nancy, 1963, cité d'après Christian Baylon, Paul Fabre, La sémantique, Paris, Nathan, 1978, p. 75

Annexe I

Généralité et centralité des propriétés



Généralité

$P_1 = 8$
 $P_2 = 5$
 $P_3 = 9$

Centralité

$P_1 = (P_1 \cap P_2) + (P_1 \cap P_3) = 3 + 4 = 7$
 $P_2 = (P_2 \cap P_1) + (P_2 \cap P_3) = 3 + 1 = 4$
 $P_3 = (P_3 \cap P_1) + (P_3 \cap P_2) = 4 + 1 = 5$

Annexe II

Propriétés	P ₁	P ₂	P ₃	P ₄	P ₅
Entités					
A	+		+		
B		+	+		+
C	+			+	+
D		+	+	+	+
E	+				
F	+				
G	+				
H	+				
I	+				
J	+				

Indices associés aux propriétés:					
	P1	P2	P3	P4	P5
Généralité:	8	2	3	2	3
Centralité:	3	5	6	5	7

Hierarchies d'intégration des entités selon:

Solidarité	Généralité	Centralité
C (10)	C (13)	D (23)
A (9)	A (11)	B (18)
E, F, G, H, I, J (7)	D (10)	C (15)
D (6)	B, E, F, G, H, I, J (8)	A (9)
B (5)		E, F, G, H, I, J (3)