

TAXINOMIE PHONOLOGIQUE ET TENDANCES UNIVERSELLES DES LANGUES DU MONDE

N. Vallée, L.J. Boë, J.L. Schwartz

*Institut de la Communication Parlée, ICP CNRS UPRESA 5009
INPG/Université Stendhal, Grenoble
(vallee, boe, schwartz)@icp.grenet.fr*

Abstract: The search for universal tendencies in the languages of the world is a necessary axis for any theoretical perspective in linguistics. We present here a typology of vowel and consonant systems from UPSID (*UCLA Phonological Segment Inventory Database*, Maddieson, 1984 ; Maddieson & Precoda, 1989). This database contains phonological data which are genetically balanced and the description of which is harmonized. We implemented it in ICP to enrich typological researches on vowels, diphthongs and consonants. We have analysed UPSID with the help of an original methodology.

We show a set of general orientations which not only confirm or make more precise some regularities already stressed on, but also brings up new data. Even if one of the conclusions is that some features are associated to specific geographical areas (the example of the languages of the African continent is significant in this matter), the phonological types and the universal tendencies brought to the fore are orthogonal to the linguistic families which were used as a source for the sampling of the database. Finally, we discuss, within the context of ontogenesis, on the general tendencies.

Keywords: typology, phonological universals, ontogenesis.

1. DE L'INTÉRÊT DES TAXINOMIES

Représentatives d'au moins 5% des langues du monde, les bases de données UPSID (*UCLA Phonological Segment Inventory Database*), d'abord 317 (Maddieson, 1984) puis 451 langues (Maddieson et Precoda, 1989), contiennent des données phonologiques génétiquement pondérées et descriptivement harmonisées. Nous les avons implantées à l'ICP pour nourrir les recherches taxinomiques sur les voyelles, les diphtongues et les consonnes.

Nous présentons ici un ensemble de tendances générales que nous exploitons comme données d'entrée pour des modèles de prédiction des structures phonologiques basés sur les propriétés des mécanismes de production et de perception des sons du langage (cf. dans ces mêmes actes "La prédiction des systèmes vocaliques : apport de la Théorie de la Dispersion-Focalisation",

J.L. Schwartz, L.J. Boë, N. Vallée & C. Abry ; aussi Schwartz & al. 1997a ; Schwartz & al. 1997b). Même s'il ressort que certains traits sont associés à certaines aires géographiques (comme c'est le cas pour les langues du continent africain), les types phonologiques et les tendances universelles mis en évidence sont orthogonaux aux familles linguistiques qui ont servi pour l'échantillonnage d'UPSID.

Nous utilisons également ces tendances générales pour tester des paradigmes d'émergence des structures sonores ; nous les discutons dans le cadre de l'ontogenèse.

2. DES UNIVERSAUX VOCALIQUES...

Les systèmes recrutent 3 à 24 voyelles, mais 80% d'entre eux en comptent moins de 10 avec une nette préférence pour un système à 5 (une langue sur 3).

Nous présentons dans la table 1b, les tendances universelles déterminées d'après UPSID₃₁₇, avec en ligne le nombre de voyelles par système et le total de langues par nombre de voyelles ; en colonne, de gauche à droite : du type le plus fréquent au moins représenté. Pour chaque type figure le pourcentage d'occurrence par nombre de voyelles. Cette répartition des systèmes les plus fréquents montre qu'il existe un type dominant par nombre de voyelles et que les systèmes concurrents sont loin derrière. Parmi ceux-ci, si on examine les plus répandus, on remarque que le système à *n* voyelles contient le système à *n*-1 et qu'il existe à l'évidence un ordre d'apparition des voyelles dans les systèmes de 3 à 9 qualités vocaliques. Aux 3 voyelles vedettes /i a u/ que l'on rencontre dans plus de 9 langues sur 10, les systèmes à 4 ajoutent de préférence la voyelle antérieure /e/, à 5 symétrisent avec la postérieure /o/. Le système /i 'e' a 'o' u/ est de loin le mieux représenté dans les langues du monde (28%). C'est le cas dans 3 des 4 grands groupes linguistiques (eurasien, américain, australien), alors que c'est le système à 7 /i e ε a o ɔ u/ qui est majoritaire en Afrique (Maddieson, 1991). Pour les systèmes pairs à 6 et 8 voyelles, les plus fréquents possèdent une voyelle centrale de type /ə/ qui pourrait renvoyer à une autre dimension liée au phénomène de la réduction vocalique (Schwartz & al., 1997a). Le type dominant à 9 contient le système à 7 auquel sont rajoutés deux timbres périphériques intermédiaires entre /i/ et /e/, entre /u/ et /o/.

Au-delà de 9 timbres de base, les systèmes, vraisemblablement saturés, utilisent une autre dimension, la nasalité en général, ou la quantité (Vallée & al., 1990 ; Vallée, 1994). Les systèmes de plus de 10 voyelles possèdent rarement plus de qualités vocaliques distinctes que les systèmes plus petits et pour augmenter leur inventaire phonologique les langues vont superposer une voire plusieurs complexités articulatoires à des timbres de base. Ainsi le système à 10 voyelles le plus fréquent possède 5 voyelles de base et 5 voyelles nasales : /i 'e' a 'o' u ɪ 'ē' ā 'ō' ũ/.

Une analyse typologique sur les dimensions de nasalité et de quantité montre que les grandes tendances mises en évidence dans les systèmes de base de 3 à 9 qualités vocaliques se retrouvent : le système à 5 /i 'e' a 'o' u/ reste le plus « populaire » et la dispersion des voyelles est périphérique et symétrique. Ces systèmes secondaires ou supplémentaires semblent relever des mêmes principes organisationnels que les systèmes de base ou systèmes primaires.

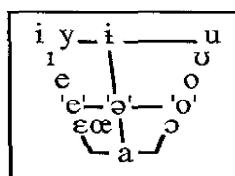


Table 1a. Qualités vocaliques de la table 1b.
























nombre de voyelles	Systèmes vocaliques et % d'occurrences (UPSID 317 langues)				
3					
20 langues	60				
4					
32 langues	37.5	18.7	6.3	6.3	
5					
134 langues	65.7	4.5			
6					
72 langues	30.6	13.9	12.5	6.9	
7					
49 langues	40.8	10.2	10.2	6.1	4
8					
34 langues	17.6	8.8	8.8	5.9	
9					
35 langues	20	14.3	6		

Table 1b. Occurrences des systèmes vocaliques par nombre de voyelles, de gauche à droite : du plus fréquent au moins répandu.
La voyelle centrale qui relève probablement d'une autre dimension est représentée par un cercle transparent.

3. ... ET CONSONANTIQUES

Les systèmes phonologiques présentent le plus fréquemment de 18 à 21 consonnes : minimum 6 pour le rotokas, langue indo-pacifique, et maximum 95 (dont 48 clicks) pour le !xū (famille khoisan). Leur répartition par mode est la suivante : occlusives 38.6%, fricatives 20.2%, nasales 14.6%, approximantes 13.0%, mi-occlusives 9.6%, vibrantes ou battues 3.9% (UPSID₄₅₁).

L'observation pour chaque langue montre que le nombre d'occlusives (orales + nasales) est beaucoup plus important que le nombre de fricatives (figure 1) à une douzaine d'exceptions près.

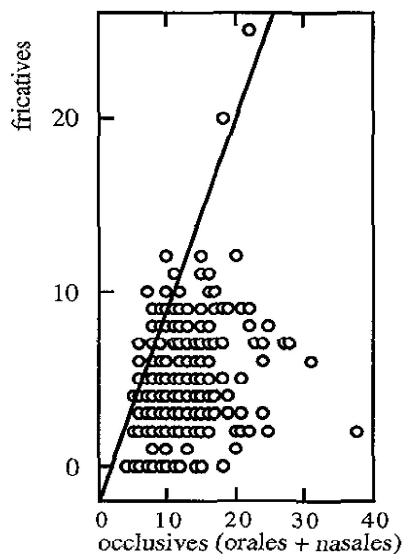


Figure 1 : Nombre de fricatives en fonction du nombre d'occlusives dans UPSID₃₁₇.

Si on comptabilise, dans UPSID₄₅₁, la distribution des occlusives par lieu d'articulation, trois classes majeures sont quasiment présentes dans tous les systèmes : coronales 100% (dentales, alvéolaires, alvéo-dentales et post-alvéolaires), bilabiales 99% et vélares 99% ; alors que : palatales 17% et uvulaires 14%. Mais on note le bon score des glottales avec 46%. Les occlusives pharyngales, qui n'étaient pas répertoriées dans UPSID₃₁₇, apparaissent dans 0.7% des langues. À l'exception des langues du continent africain, les sourdes dominent largement les sonores (64%) : on relève 30% de sourdes en plus dans les bilabiales, 50% de sourdes en plus dans les coronales, 60% dans la catégorie vélaire.

Ce rapport sourdes/sonores est encore plus écrasant pour les fricatives : les langues présentent 5 fois plus de /s/ que de /z/, 3,2 fois plus de /ʃ/ que de /ʒ/ et presque 2 fois plus de /f/ que de /v/. Quel que soit le lieu d'articulation, le trait sourd concerne 72% des fricatives. Les mi-occlusives sont également majoritairement sourdes (74%).

La table 2 présente, par ordre décroissant, les pourcentages d'occurrences des 25 consonnes les plus fréquentes relevées dans UPSID₄₅₁.

Table 2. Phonèmes consonantiques les plus fréquents
/t d n l/ regroupent toutes les coronales (UPSID₄₅₁).

C	%	C	%
t	97.5	g	56.2
m	94.4	ŋ	52.7
n	90.4	ʔ	48.0
k	89.5	tʃ	41.8
j	84.0	ʃ	41.6
p	83.3	f	40.0
w	76.8	dz	34.9
s	73.5	ɲ	31.3
d	64.7	ts	29.3
b	63.8	kʰ	22.9
h	62.0	pʰ	22.4
l	56.9	v r	21.1

4. CONSONNES VS. VOYELLES

Pour l'ensemble des langues la corrélation entre le nombre de consonnes et le nombre de voyelles n'est pas significative (figure 2). Seule régularité, chaque langue possède plus de consonnes que de voyelles (à 2 exceptions près : le pawaïan (famille austro-thaï) avec 12 voyelles et 10 consonnes ; et l'apinaye (famille sud-amérindienne, groupe macro-ge), 17 voyelles et 13 consonnes).

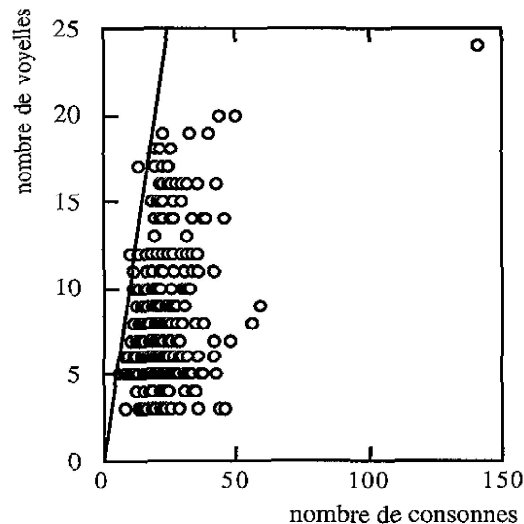


Figure 2 : Nombre de voyelles en fonction du nombre de consonnes dans les systèmes d'UPSID₃₁₇.

5. UNIVERSAUX ET ONTOGENÈSE

La revue de question de Locke (1983) rend compte de l'inventaire des consonnes représentant 95% des occurrences produites par des enfants âgés de 11 à 12 mois, au stade du babillage canonique ou panaché. Les données portent sur 129 sujets pour 15 langues et livrent, par ordre décroissant de fréquence : /b/ et /m/, /p/, /d/, /h/ et /n/, /t/, /g/ et /k/, /j/ et /w/, /s/. On ne dispose pas d'un tel inventaire aussi précis pour les voyelles ; nous savons qu'elles sont beaucoup moins nombreuses que les consonnes.

On constate alors que ces 12 consonnes se retrouvent exactement – au /l/ près qui est acquis plus tardivement – parmi les consonnes les plus fréquentes dans l'inventaire des langues d'UPSID, tel que cela apparaît dans la table 2.

Dans ce peloton de tête les *occlusives* (*nasales* comprises) dominent nettement et les *fricatives* ne sont représentées que par /s/ (et /h/, classé fricatif par Maddieson comme par l'API, mais dont le statut est pour le moins spécifique, cf. le nœud laryngal dans la géométrie des traits). Les fricatives sont donc bien les plus rares dans le babillage et c'est aussi ce que l'on observe dans la base des langues du monde. À partir de cette tendance – parmi d'autres –, Locke et Pearson (1992, p. 96) n'hésitent pas à tirer, à juste titre, la conclusion de portée générale suivante :

« Infants heavily favor stop consonants over fricatives, and there are languages that have stops and no fricatives but no languages that exemplify the reverse pattern. [Such] "phonologically universal" patterns, which cut across languages and speakers are, in effect, the phonetic properties of *Homo sapiens*. »

6. CONCLUSIONS ET PERSPECTIVES

La mise en parallèle des tendances universelles des systèmes sonores des langues du monde avec les données concernant l'acquisition du langage nous permet de formuler l'hypothèse que les sons des langues sont vraisemblablement puisés dans les stocks des capacités potentielles du babillage.

Corroborant cette hypothèse, les typologies phonologiques et les tendances universelles mises en évidence montrent que classement *typologique* et classement *génétique* sont loin d'être identiques, et que la typologie a des tendances *géographiques*. Cet état des choses a fait l'objet de nombreuses remarques depuis Jakobson (1936) ; cf. Ruhlen (1978) pour les nasales.

Une des inférences intéressantes de ce manque de relations est que les tendances générales des systèmes pourraient dépendre des capacités sensori-motrices de l'*Homo loquens* ; et il semble qu'on puisse en trouver une confirmation dans l'*ontogenèse*, appuyant ainsi la thèse qu'existeraient des « propriétés phonétiques d'*Homo sapiens* ».

7. RÉFÉRENCES

- Jakobson R. (1936) Sur les affinités phonologiques entre les langues. *4^e Congrès International de Linguistes*, 48-58.
- Locke J.L. (1983) *Phonological acquisition and change*. Academic Press, New-York.
- Locke J.L., Pearson D.M. (1992) Vocal Learning and the Emergence of Phonological Capacity. A Neurobiological Approach. In *Phonological Development. Models, Research, Implications*. C. Ferguson, L. Menn & C. Stoel-Gammon. (Eds.) Timonium, Maryland, 91-129.
- Maddieson I. (1984). *Patterns of sounds*. Cambridge University Press, Cambridge (2^e édition 1986).
- Maddieson I. (1991) Testing the Universality of Phonological Generalizations with a Phonetically Specified Segment Database: Results and Limitations. *Phonetica* 48, 193-206.
- Maddieson I. & Precoda K. (1989) Updating UPSID. *UCLA WPP* 74, 104-111.
- Ruhlen M. (1978) Nasal vowels. In *Universal of Human Language*. J. Greenberg & al. (Eds.), vol. 1, 203-241.
- Schwartz J.L., Boë L.J., Vallée N., & Abry C. (1997a) Major trends in vowel system inventories. *J. of Phonetics* 25, Vol. 3, 233-253.
- Schwartz J.L., Boë L.J., Vallée N., & Abry C. (1997b) The Dispersion-Focalization Theory of vowel systems. *J. of Phonetics* 25, Vol. 3, 255-286.
- Schwartz J.L., Boë L.J., Vallée N., & Abry C. (1997) La prédiction des systèmes vocaliques : apport de la théorie de la dispersion-focalisation. *16^e CIL*, dans ces mêmes Actes.
- Vallée N. (1994) *Systèmes vocaliques : de la typologie aux prédictions*. Thèse de Doctorat en Sciences du Langage, Université Stendhal, Grenoble.
- Vallée N., Boë L.J., & Schwartz J.L. (1990) Systèmes vocaliques : typologie et tendances universelles. *18^e Journées d'Étude sur la Parole*, 32-36.