



Linx

Revue des linguistes de l'université Paris X Nanterre

45 | 2001

Invariants et variables dans les langues. Études typologiques

Des lexiques aux syllabes des langues du monde

Typologies, tendances et organisations structurelles

Nathalie Vallée, Isabelle Rousset et Louis-Jean Boë



Édition électronique

URL : <http://journals.openedition.org/linx/721>

DOI : 10.4000/linx.721

ISSN : 2118-9692

Éditeur

Presses universitaires de Paris Nanterre

Édition imprimée

Date de publication : 1 décembre 2001

Pagination : 37-50

ISSN : 0246-8743

Référence électronique

Nathalie Vallée, Isabelle Rousset et Louis-Jean Boë, « Des lexiques aux syllabes des langues du monde », *Linx* [En ligne], 45 | 2001, mis en ligne le 19 juin 2012, consulté le 19 avril 2019. URL : <http://journals.openedition.org/linx/721> ; DOI : 10.4000/linx.721

Ce document a été généré automatiquement le 19 avril 2019.

Département de Sciences du langage, Université Paris Ouest

Des lexiques aux syllabes des langues du monde

Typologies, tendances et organisations structurelles

Nathalie Vallée, Isabelle Rousset et Louis-Jean Boë

NOTE DE L'AUTEUR

Nos remerciements à Ian Maddieson pour les lexiques, Delphine Zouyed, Céline Bard pour leur contribution à l'implantation d'ULSID. Une partie de cette recherche a été financée par le GDR-CNRS 1954 "Phonologie".

1. Introduction

- 1 Toute unité lexicale, quelle que soit la langue, est constituée d'un ou plusieurs types syllabiques existant dans cette langue. Une structure syllabique correspond à une concaténation de phonèmes décomposable traditionnellement en constituants C et V. Cependant, la nature de l'unité syllabique est encore actuellement l'objet d'un débat linguistique. Beaucoup on fait l'hypothèse que l'existence de l'unité syllabique ne pourrait être montrée que par l'existence de ses frontières. Des études pour délimiter la syllabe ont surtout été menées dans l'exploration des limites des phénomènes de coarticulation, dans les processus de lénition ou de fortition (*cf.* l'état de l'art de Krakow, 1999). Mais ces recherches, comme les études acoustiques d'ailleurs, n'ont pas livré à ce jour de réponse satisfaisante et complète sur l'unité syllabique. Depuis une dizaine d'années, les études sur la syllabe se penchent plutôt sur la recherche de caractéristiques articulatoires qui seraient plus spécifiques des attaques syllabiques et d'autres qui seraient des indices révélateurs de finales syllabiques. Dans le même ordre d'idée, les études menées sur les asymétries attaque/coda soutiennent le rôle organisationnel de l'élément syllabique pour les modèles phonologiques et visent à proposer une définition

claire et précise de l'unité syllabique en identifiant la nature des représentations phonologiques.

- 2 Sans disposer d'une définition de la syllabe cohérente à plusieurs niveaux de description linguistique, la recherche et l'analyse des séquences généralement favorisées ou défavorisées dans les langues du monde, des fréquences des combinaisons inter- et intra-langue au niveau syllabique et lexical, des dépendances et indépendances distributionnelles des unités phoniques inter- et intra-syllabiques, montrent qu'il est tout à fait légitime de penser que l'organisation syllabique joue un rôle dans la structuration de la chaîne parlée et dans la formation des unités lexicales. Les tendances donnent de fortes raisons pour que l'organisation syllabique ne soit pas à considérer comme le fruit du hasard. Mais établir une typologie des langues basée sur la syllabe, en regroupant les structures syllabiques présentant des similitudes phoniques ou organisationnelles, est insuffisant pour élaborer des concepts phonologiques. Les typologies ne comportent pas en elles mêmes de pouvoir explicatif : elles aident à formuler des hypothèses qu'elles ne peuvent pas à elles seules prétendre vérifier. Aujourd'hui, il apparaît de plus en plus difficile, voire impossible, de tenter d'expliquer les tendances générales des structures sonores des langues du monde avec des principes internes de phonologie. Pour éviter le "*curve-fitting*" qui priverait la typologie linguistique d'explication scientifique (Lindblom, 2000), les processus linguistiques ne peuvent pas être tirés des données observables qui doivent être expliquées par ces processus. Depuis une trentaine d'années, les explications des tendances phonologiques des langues du monde, quelles que soient les familles linguistiques, sont recherchées dans les aspects articulatoires, acoustiques et psychologiques de la parole (cf. par exemple, Lindblom, 2000 Lindblom *et al.*, 1984, Schwartz *et al.*, 1997 ; Vallée *et al.*, 1999).
- 3 Une typologie des structures syllabiques est fondamentale si l'on veut faire le point sur des discussions de questions générales de phonologie : peut-on établir une organisation universelle des structures syllabiques ? Peut-on dresser des inventaires de différents types de structures syllabiques selon les langues, selon les segments ? Dans quelle proportion les séquences associant consonnes et voyelles peuvent être déterminées par des caractéristiques articulatoires et perceptives des sons impliqués dans la cohorte ? Quelles sont les restrictions qui opèrent ? La différenciation des propriétés acoustiques et articulatoires des constituants de la syllabe est-elle déterminante dans l'organisation syllabique ? La distinctivité joue-t-elle un rôle dans l'émergence du lexique ?
- 4 Dans notre propos, il ne s'agit pas d'analyser le statut linguistique de la syllabe, mais son contenu, afin de mettre en évidence les facteurs susceptibles de conditionner des choix ou des restrictions en faisant appel à la substance des unités sonores en contexte immédiat ou proches. Ces investigations dans le domaine de la syllabe visent à prolonger les connaissances que nous avons déjà acquises sur les contraintes biologiques qui structurent en partie les systèmes phonologiques des langues naturelles (Vallée *et al.*, 1999 ; Boë *et al.*, 2000).
- 5 Dans l'état actuel des avancements de nos travaux, nous nous limitons ici à un exposé de résultats préliminaires dans lequel figure un certain nombre d'observations qui permettent déjà de dégager des caractéristiques typologiques basées sur la syllabe et de préciser des considérations générales sur les types syllabiques de 13 langues.

2. ULSID

- 6 Ian Maddieson a mis à notre disposition une banque de données contenant les lexiques de 32 langues découpés en syllabes (Maddieson, 1992). Baptisée ULSID, acronyme de *ucla - université de los angeles - lexical and syllabic inventory database*, elle contient des langues sélectionnées dans un souci de diversité à la fois génétique et géographique : hétérogénéité des familles de langues et très large répartition géographique. Les langues retenues disposent toutes d'un dictionnaire ou d'un lexique dont les entrées sont soit phonétiques, soit phonologiques, soit orthographiques lorsque le code graphique de la langue est aisément interprétable avec un code phonétique. La totalité des lexiques est alors transcrite en ASCII avec deux types de séparateurs : l'un indiquant le découpage de chaque entrée en syllabe, l'autre codant la séparation entre attaque, noyau et coda de chaque syllabe. Nous donnons à titre d'exemple, trois entrées en ASCII du lexique du kwakw'ala (langue almosan), 2 des items étant dissyllabiques, un troisième trisyllabique :

```
k u Xw s .? a
k u Xw .ts' a .n a
k u Xw .? i d
```

- 7 Dans l'étape d'implantation des données d'ULSID à l'ICP, nous avons harmonisé la transcription entre les lexiques, en adoptant les symboles conventionnels de l'API (1996) et en conservant l'ensemble des informations sur le découpage syllabique. Actuellement, 13 langues sur 32 ont été transcodées en retournant aux sources phonétiques et phonologiques de chaque langue. Elles présentent une diversité dans la taille et l'inventaire de leurs systèmes vocaliques et consonantiques. C'est de cet échantillon (cf. Table 1), resté représentatif de la diversité mentionnée plus haut, que nous avons tiré les premiers résultats.

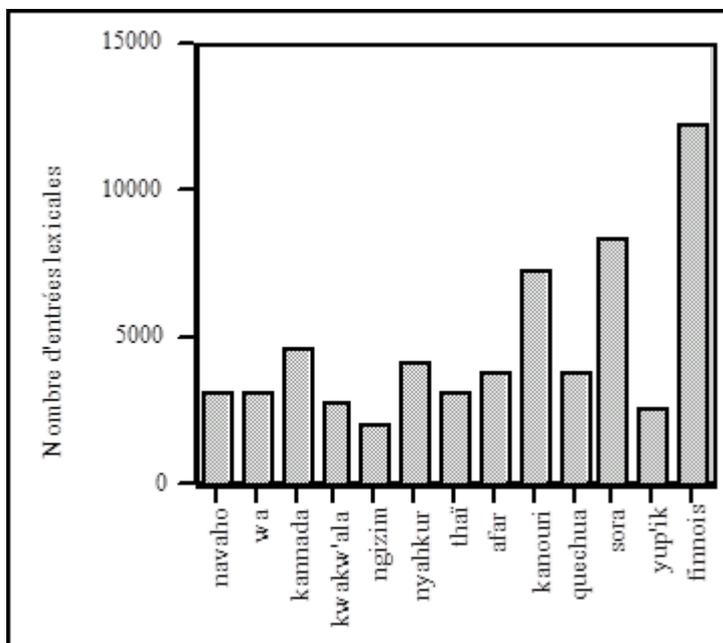
Table 1. Les 13 langues d'ULSID

langue	famille, branche	lieu géographique
wa	austro-asiatique, palaung	Chine
kannada	dravidien, sud-dravidien	Inde : état du Mysore
sora	austro-asiatique, mounda	Inde : états d'Orissa et d'Assam
thaï	kam-taï, kadaï,	Thaïlande, Birmanie, province du Yunnan
nyahkur	austro-asiatique, môn	Thaïlande (centre)
ngizim	afro-asiatique, tchadique	Nord-est du Nigéria
afar	afro-asiatique, couchitique	corne orientale de l'Afrique
kanouri	nilo-saharien, saharien	Nigéria, Niger, Tchad, Cameroun
navaho	na-déné, athabaskan	USA : Arizona, Colorado, Nouveau-Mexique, Utah

kwakw'ala	Amérique du Nord, almosan,	Canada : Colombie Britannique
quechua	Amérique du Sud, Andin	Pérou, Équateur, Bolivie, Colombie
yup'ik	eskimo-aléoute, eskimo	Alaska
finnois	ouralo-altaïque, finno-ougrien	Finlande

- 8 La taille des lexiques est portée sur la figure 1. Au total, nous disposons d'un peu plus de 160 500 syllabes résultant du découpage de 60 000 entrées lexicales, avec une moyenne de 4 560 termes par langue. Les emprunts récents, dont les référents sont d'ordre technologique, politique ou culturel, ont été soustraits des lexiques par Maddieson (1992). Les informations concernant le découpage syllabique ont été livrées soit par les dictionnaires, soit à partir d'informateurs, locuteurs natifs des langues concernées.

Figure 1. Taille des lexiques

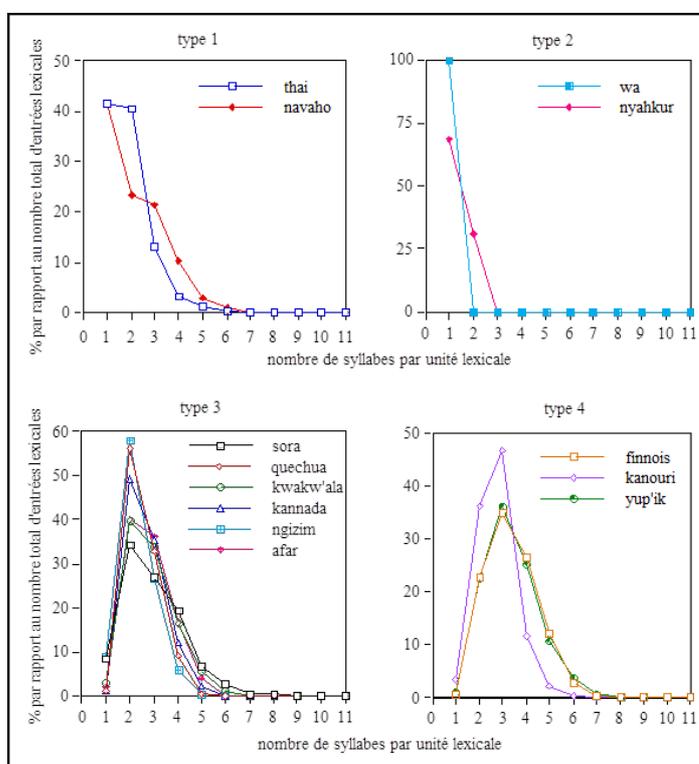


3. Structure syllabique des unités lexicales et typologie des langues

- 9 Ne disposant d'aucune information sur la nature et le sens des unités lexicales, elles sont toutes prises en compte dans les résultats.
- 10 La répartition des lexiques de chaque langue en fonction du nombre de syllabes contenues dans les entrées lexicales, permet de dégager 4 types (Figure 2) :
- Le type 1 regroupe les langues qui présentent une distribution des unités lexicales telle que 40 % au moins d'entre elles sont monosyllabiques, entre 20 et 40 % sont dissyllabiques, entre 10 et 20 % sont trisyllabiques (navaho, thai) ;

- Le type 2 rassemble les langues totalement ou très majoritairement monosyllabiques : le wa (100 %), le nyahkur (70 % d'unités monosyllabiques et 30 % de dissyllabes). Notons que ces langues, contrairement à celles du type 1, ne possèdent pas d'items lexicaux de plus de 2 syllabes ;
 - Sous le type 3 se classent les langues qui présentent une majorité d'items dissyllabiques, avec moins de 10 % d'unités monosyllabiques et environ un tiers d'entrées lexicales trisyllabiques. Il s'agit du type de répartition le plus répandu dans les langues de notre échantillon ;
 - Le type 4 regroupe le finnois, le kanouri et le yup'ik, qui ne présentent que très peu d'unités monosyllabiques dans leur lexique, mais une majorité d'unités trisyllabiques, et 20 à 40 % de dissyllabes, 25 % de quadrisyllabes, de 2 (kanouri) à 12 % (finnois) de quinquasyllabes.
- 11 Sous chaque type figurent des langues éloignées géographiquement et de parenté linguistique différente, sauf pour le type 2 qui regroupe deux langues austro-asiatiques mais appartenant à deux branches différentes : môn pour le nyahkur et palaung pour le wa (cf. Table 1).

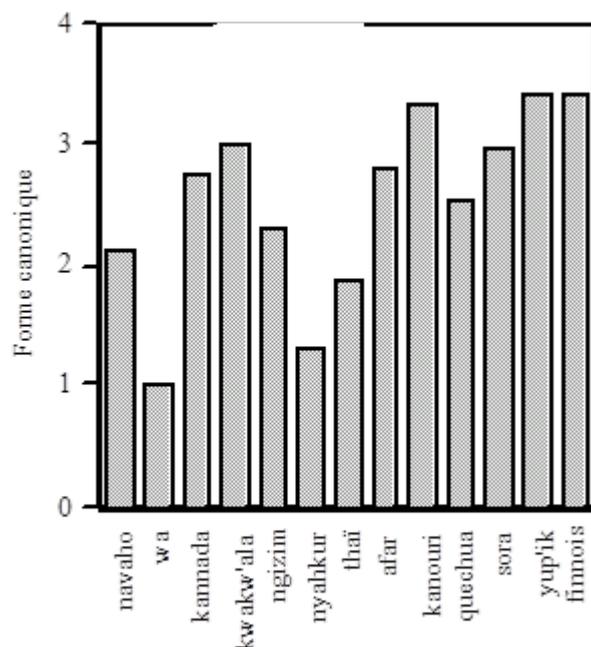
Figure 2. Typologie des 13 langues d'ULSID, élaborée sur la répartition du lexique en fonction du nombre de syllabes par item lexical



- 12 De toute évidence, ce sont les structures lexicales di- et trisyllabiques qui sont très largement favorisées dans les lexiques des langues de notre échantillon, mises à part, bien entendu, celles relevant du type 2. Toutefois, 8 langues sur 13 favorisent les dissyllabes, contre 3 privilégiant les unités trisyllabiques (Figure 3). Même si deux des trois langues du type 4 possèdent les lexiques les plus importants d'ULSID, les unités lexicales dissyllabiques restent majoritaires dans le regroupement des 13 lexiques. Notre typologie montre que les éléments lexicaux de plus de 5 syllabes sont marginaux (entre 0 et moins de 4 % du lexique des langues), ceux à 5 syllabes constituant au maximum un peu plus de

10 % du lexique pour le finnois et le yup'ik (de type 4) et entre 0 et moins de 6 % du lexique pour les autres langues.

Figure 3. Forme canonique (Fc) des langues selon la formulation de Monino et Roulon (1972).

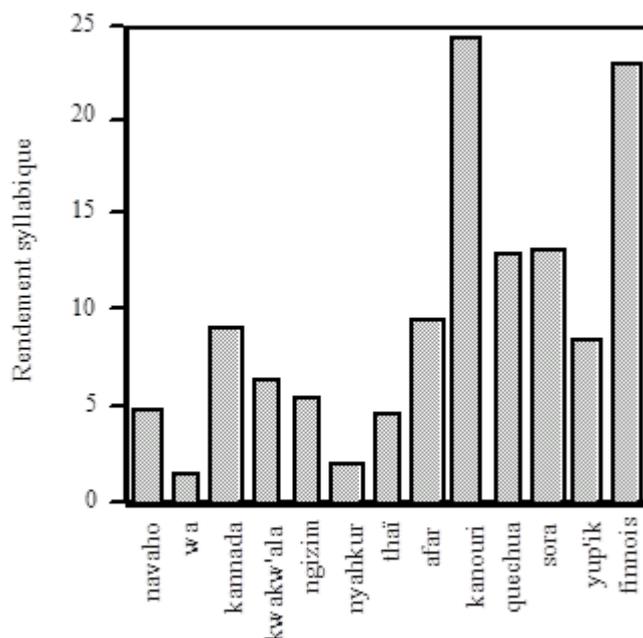


Fc correspond au rapport entre le nombre total de syllabes pour l'ensemble d'un lexique donné et le nombre d'entrées lexicales. Les valeurs obtenues pour Fc montrent bien la préférence des langues pour les formes lexicales di- et trisyllabiques.

4. Rendement syllabique

- 13 Pour les lexiques de moins de 5 000 termes, on ne relève pas de corrélation très nette entre la taille du lexique et le nombre total de syllabes qui le constituent. Seules trois langues (sora, kanouri, finnois) présentent entre 7 000 et 12 000 entrées lexicales et un nombre de syllabes allant de 19 700 à plus de 41 000. Pour ces 3 langues, on note un large nombre de syllabes lié à la taille plus importante du lexique.
- 14 Sans faire d'équivalence typologique sur la nature des segments qui composent les syllabes, nous avons procédé à un regroupement des syllabes identiques pour chacune des langues. Nous avons déterminé le rendement syllabique comme étant le rapport entre le nombre total de syllabes obtenues dans le découpage d'un lexique donné et le nombre de syllabes différentes comptabilisées après regroupement (Figure 4). Le rendement syllabique permet de rendre compte de la fréquence de l'ensemble des syllabes d'une langue dans les items lexicaux.

Figure 4. Rendement syllabique pour chaque langue, calculé comme le rapport du nombre total de syllabes et du nombre de syllabes différentes après regroupement des syllabes identiques.



- 15 La tendance à observer un rendement syllabique caractéristique pour chacun des 4 types de notre typologie concerne 12 langues sur 13. On constate un rendement élevé pour le finnois ou le kanouri qui sont des langues relevant du type 4, et qui se caractérisent, nous l'avons vu, par une forte proportion d'unités lexicales trisyllabiques. Les rendements les plus faibles sont rencontrés pour les langues du type 2, c'est-à-dire, celles totalement ou très majoritairement monosyllabiques. Un rendement proche de 1 signifie que la fréquence des syllabes dans le lexique de la langue concernée est relativement faible. Un rendement proche de 5 caractérise les langues du type 1 (navaho et thaï), alors que les langues du type dominant (type 3) ont un rendement syllabique allant de plus de 5 à 14.

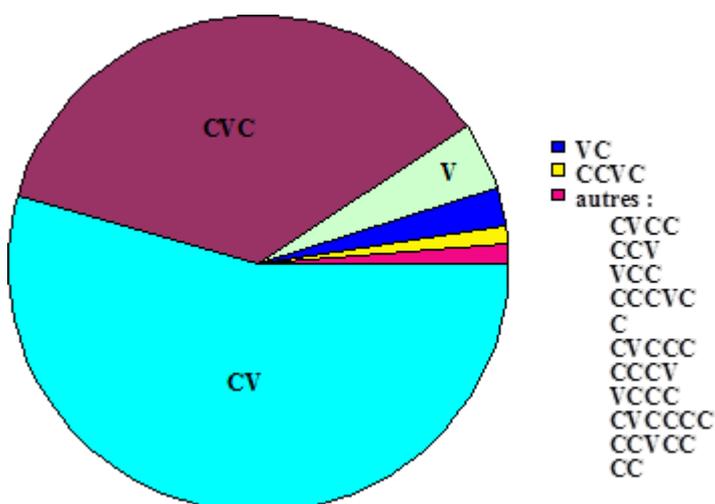
5. Typologie et tendances des structures syllabiques

- 16 La décomposition des syllabes de chaque lexique en constituants C et V, et leur regroupement en structure identique, révèlent que le nombre de type est relativement restreint, quelle que soit la langue, et qu'il varie dans un intervalle allant de 4 à 11, avec une moyenne de 7.7 types par langue. Le contingent de types de syllabes pour chaque langue est totalement indépendant de la taille du lexique, du nombre total de syllabe et du rendement : citons le kannada, le nyahkur et le kanouri qui possèdent chacune 7 types de structures syllabiques avec respectivement 4 559, 4 188, 7 211 entrées lexicales, et 12 126, 5 503, 19 773 syllabes au total, ainsi que des rendements de 9.12 pour le kannada, 1.88 pour le nyahkur et 24.23 pour le kanouri. Cette typologie de la structure syllabique appuie l'existence de restrictions qui organisent la syllabe dans les langues naturelles (Table 2) : les 160 517 syllabes, totalisées sur l'ensemble de notre corpus, se répartissent dans 16 types de structures syllabiques et plus de 90 % d'entre elles sont de type CV ou CVC (Figure 5).

Table 2. Liste et occurrence des structures syllabiques rencontrées dans les 13 langues d'ULSID

type	0/000	type	0/000
CV	5448	CCCVC	2
CVC	3615	C	1.4
V	440	CVCCC	1.4
VC	247	CCCV	0.3
CCVC	126	VCCC	0.3
CVCC	63	CVCCCC	0.2
CCV	50	CCVCC	0.2
VCC	5	CC	0.2

Figure 5. Fréquence de chaque structure syllabique dans les 13 langues.



- 17 Les syllabes fermées présentent plus de diversité : 11 types contre 5 pour les syllabes ouvertes. Les deux structures que l'on rencontre dans toutes les langues de la base sont CV et CVC, mais pour 10 d'entre elles, c'est la première qui est la plus répandue, le wa, le nyahkur et le thaï (toutes trois issues de la famille austro-asiatique mais de branches différentes) possédant majoritairement des syllabes CVC (Figures 6 et 7). De manière plus générale, ces trois lexiques (regroupés sous les types 1 et 2 de notre typologie des langues) favorisent, contrairement aux autres, les syllabes fermées. On note également que 8 langues sur 10 comptent 60 % ou plus de syllabes dans le type CV. De ce fait, la tendance très forte qui se dégage de cette typologie est que les groupements consonantiques intra-syllabiques sont nettement défavorisés : on ne les rencontre que dans 1.26 % des quelques 160 500 syllabes de notre corpus. Si on se penche sur la place de ces groupements consonantiques dans la syllabe, nous constatons que la plupart (67 %)

occupent la position initiale, contre 33 % en finale de syllabe. Ce résultat est renforcé par le fait que les consonnes complexes (celles qui superposent à une articulation de base des modes articulatoires tels que la labialisation, l'aspiration, la glottalisation, la palatalisation, la prénasalisation... mais qui n'occupe qu'une position C dans notre corpus) sont bien plus fréquentes en attaque de syllabe. Encore nous reste-t-il à déterminer des tendances générales dans la nature des consonnes qui apparaissent dans les groupements.

Figure 6. Répartition des types de syllabes ouvertes par langue (seuls figurent les plus fréquents).

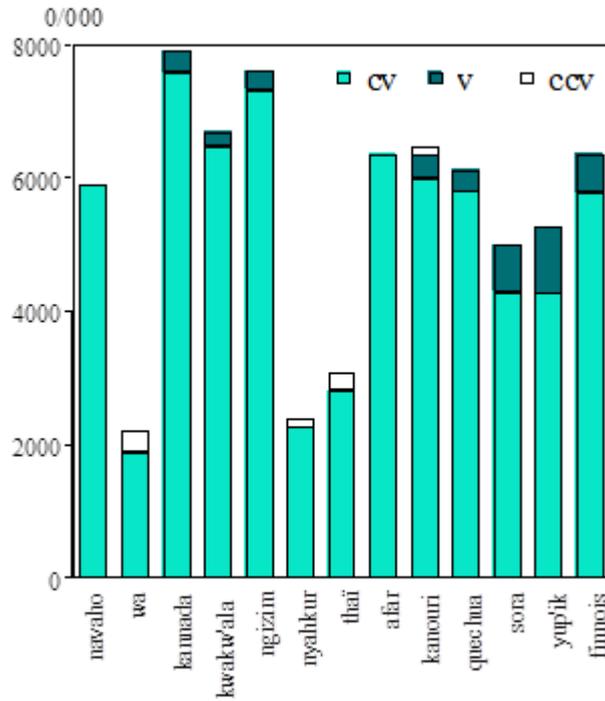
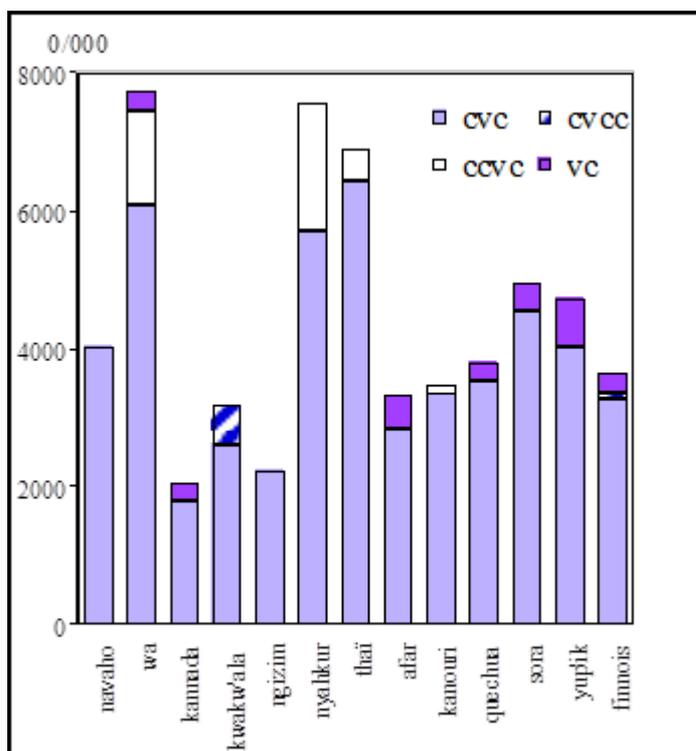


Figure 7. Répartition par langue des types de syllabes fermées les plus répandus.



- 18 Un ensemble de tendances implicationnelles émerge de l'observation de la fréquence des 16 structures syllabiques dans nos 13 langues. En effet, sans tenir compte des caractéristiques des consonnes impliquées dans les groupements, on constate que la complexité des structures syllabiques ouvertes ou fermées d'une langue implique la présence de la même structure (ouverte ou fermée) moins complexe ; le degré de complexité se mesurant ici par le nombre de segments constituant l'attaque ou la coda.
- 19 Ainsi, on observe que si une structure syllabique avec attaque complexe de n consonnes est attestée dans une langue, elle implique les structures avec attaque de $n - i$ consonnes ($i = 1$ à $n - 1$) :

ccccv	→	cccv	→	ccv	→	cv
cccvc	→	ccvc	→	cvc		

- 20 La même tendance est observée pour la complexité des codas, ce qui signifie que si une structure syllabique avec coda complexe de n consonnes est attestée dans une langue, elle implique les structures avec coda de $n - i$ consonnes ($i = 1$ à $n - 1$) :

cvcccc	→	cvccc	→	cvcc	→	cvc
vccc	→	vcc	→	vc		

- 21 De manière très régulière, la complexité est inversement proportionnelle à la fréquence d'occurrence. En effet, si le nombre de segments n présents en attaque ou en coda augmente, le rang d'occurrence de la structure syllabique diminue :

CCCCV	<	CCCV	<	CCV	<	CV
CCVC	<	CCVC	<	CVC		
CVCCCC	<	CVCCC	<	CVCC	<	CVC

- 22 La fréquence d'une structure syllabique est donc liée au degré de complexité de l'attaque et/ou de la coda, celui-ci correspondant au nombre de segments occupant cette position.
- 23 Les mêmes tendances sont observables dans les structures à attaque vide. En effet, les structures syllabiques avec attaque vide impliquent la présence de la structure V dans la langue selon les implications suivantes :

vccc	→	vcc	→	vc	→	v
-------------	---	------------	---	-----------	---	----------

- 24 et la fréquence des structures avec attaque vide est aussi liée à la complexité de la coda (ne sont relevées que deux exceptions dans des proportions qui restent faibles en wa et en afar, cf. Table 3) :

vccc	<	vcc	<	vc	<	v
-------------	---	------------	---	-----------	---	----------

Table 3. Proportion des structures V et VC dans les syllabes de l'afar et du wa.

	afar	wa
v	3.15 %	0.7 %
vc	4.65 %	2.74 %

- 25 Remarquons que ces tendances confirment et prolongent celles de Blevins (1996 : 217-220) puisque l'échantillon qu'elle a retenu pour mener son étude est constitué des structures syllabiques de 24 langues dont 1 seule, le finnois, figure dans les 13 langues d'ULSID.
- 26 La comparaison des fréquences d'occurrence entre les structures ouvertes et fermées montrent que la préférence des langues pour un type syllabique ouvert ou fermé dans les structures de base CV ou CVC est conservée dans les structures plus complexes. En effet, à degré de complexité équivalent, si une langue favorise les structures ouvertes alors les structures ouvertes à attaque complexe (type CCV) sont plus répandues dans cette langue que les structures fermées à attaque complexe (type CCVC). Inversement, si une langue favorise les structures fermées, alors les structures à coda complexe seront plus fréquentes dans cette langue que les structures à attaque complexe (CVCC vs. CCVC).
- 27 Cependant, si les langues favorisent de très loin les attaques et codas simples dans leurs structures syllabiques, les articulations vocaliques ou consonantiques élaborées peuvent être rencontrées parmi les syllabes les plus fréquentes dans une langue donnée (exemples [ka:n] [t^h i] ou [K^hwa:m] qui sont respectivement les syllabes de rang 2, 4 et 5 du thaï). La comparaison des 20 syllabes les plus fréquentes pour chaque langue, montre qu'elles possèdent majoritairement une attaque constituée de plosives sourdes, vélaires et

coronales [k t], devant les latérales coronales [l], les nasales coronales et bilabiales [m n] et la fricative coronale sourde [s]. Les noyaux de ces syllabes fréquentes sont pour plus d'une syllabe sur deux occupés par la voyelle ouverte [a], devant [i] et [u]. Nous remarquons également que [a] constitue le noyau le plus répandu des syllabes qui présentent une attaque et une coda vide (type V). Lorsqu'une langue n'est pas en adéquation avec cette tendance, c'est que la structure V est marginale ou inexistante sur l'ensemble du lexique (navaho, wa, nyakhur, thaï). Il est plus difficile de tirer une observation générale sur la nature des consonnes en coda de syllabes fréquentes dans une langue car pour beaucoup c'est la structure CV qui prédomine. À noter une forte proportion de [n] dans cette position. Parmi les syllabes les plus répandues dans une langue donnée, l'inventaire des consonnes en coda est bien plus restreint que l'inventaire des possibilités pour les consonnes en attaque, excepté pour une langue de notre échantillon, l'afar (cf. Table 4).

Table 4. Taux de l'inventaire consonantique pouvant apparaître en attaque ou en coda pour chaque langue

	attaque	coda
afar	100 %	100 %
		94 %
finnois	100 %	82,35 %
		50 %
kannada	100 %	92,6 %
		70 %
kanouri	95,5 %	68,2 %
		41 %
kwakw'ala	100 %	59,5 %
		33 %
navaho	100 %	37,8 %
		27 %
ngizim	97,3 %	70 %
		38 %
nyakhur	86,7 %	62 %
		52 %
quechua	97 %	66,66 %
		30 %
sora	92 %	54 %
		26 %
thaï	95,5 %	50 %
		41 %
wa	97,3 %	21 %
yup'ik	92,7 %	65,9 %
		27 %

En grisé, la part des consonnes apparaissant dans plus de 1 % des codas. Sont exclues des rapports les structures syllabiques avec groupement consonantique.

6. Discussion

- 28 Notre étude typologique avec recherche de grandes tendances dans le formalisme des structures syllabiques de 13 langues a pour but, dans un premier temps, de mettre en lumière une organisation intra-syllabique dans les cohortes de segments. La présence de formes majoritairement favorisées dans les langues est un élément fondamental qui s'impose dans les discussions générales de phonologie.

- 29 Des travaux récents en phonétique - phonologie et en acquisition du langage montrent que l'étude de la syllabe est devenue un cadre nécessaire et incontournable pour comprendre le fonctionnement du langage et apporter des éléments intéressants de discussion à des questions et des problèmes d'ordre général non encore résolus : la syllabe est-elle l'unité fondamentale du langage humain ? Quels éléments pertinents pourraient appuyer l'existence phonétique, phonologique et psychologique de l'unité syllabique ? Dans quelles proportions le matériau phonologique universel des langues est-il basé sur des potentialités articulatoires reliables à la syllabe ? Comment éviter les hypothèses implicites sur l'organisation syllabique dans les modèles phonologiques ? Quels constituants déterminants de la syllabe faut-il prendre en compte dans une théorie générale de la syllabe ?
- 30 Si la syllabe est l'unité de base de la parole (MacNeilage, 1998 ; MacNeilage et Davis, 2000), on doit pouvoir trouver les traces d'une organisation syllabique dans les langues du monde. Il s'agit pour nous de récupérer, par l'observation des structures syllabiques des lexiques de langues naturelles - à la différence de Janson (1986) qui a considéré la fréquence des syllabes dans des textes de 4 langues - les informations nécessaires qui nous permettront de montrer que si des formes de syllabes sont nettement favorisées dans les langues c'est qu'elles sont probablement plus fonctionnelles que d'autres. Bien sûr pour arriver à monter l'existence de contraintes fonctionnelles dans les structures syllabiques, nos investigations sont pour l'heure largement insuffisantes. N'ont pas encore été étudiées la fréquence des syllabes sur l'ensemble du corpus, les contraintes sur les cohortes, les associations consonantiques et consonne-voyelle en fonction de la position dans la syllabe (attaque, coda), ainsi que la nature de la voyelle en fonction du lieu d'articulation de la consonne et ceci, en nous appuyant sur des résultats typologiques que nous avons déjà acquis (Vallée *et al.*, 1999 ; Vallée et Boë, 2000).
- 31 Nous faisons l'hypothèse que la fréquence importante de la structure CV, la marginalisation des groupements consonantiques intra-syllabique et la forte proportion d'unités lexicales dissyllabiques dans les langues pourrait très bien être liée, non pas à une exigence formelle, mais à des contraintes de production. Notre objectif est de tenter d'analyser les points cités ci-dessus en considérant le geste articulatoire de fermeture et d'ouverture du tractus vocal rythmé par la mâchoire, selon la théorie *frame/content* de MacNeilage (1998). Dégager les schèmes structurels des différents lexiques est une étape qui s'imposera dans cette recherche, de même que confronter les tendances des structures syllabiques aux données de l'ontogenèse (Vihman, 1996) puisque les cohortes CV sont aussi les syllabes canoniques du babillage de l'enfant, quel que soit son environnement linguistique (Locke, 1983). Ces tendances peuvent aussi étayer l'existence d'un attracteur CV au détriment des autres types de syllabe, et de VC en particulier, qui fait converger les gestes syllabiques vers une forme plus stable et mieux résistante aux contraintes de production. Cet attracteur est mis en évidence avec un paradigme de contrainte de production basé sur le débit, le locuteur devant produire des syllabes avec un débit de plus en plus accéléré, synchronisé sur un métronome (Tuller et Kelso, 1991). Désorganisant le contrôle de synchronisation supposé coûteux, les syllabes VC « *switchent* » vers l'attracteur CV. Par ailleurs, différentes études (cf. Krakow, 1999) mettent en évidence qu'en attaque de syllabe, la force articulatoire est plus grande, le geste plus précis, et que les segments montrent une plus grande stabilité, une résistance plus importante à la coarticulation. De tels avantages expliqueraient que les langues favorisent cette structure.

- 32 Dans l'esprit de ce qui est plaidé par Lindblom (2000), Lindblom *et al.* (1984) à savoir, l'optimisation des séquences de segments à partir des propriétés des systèmes humains de production et de perception de parole, nous voulons, à terme, apporter des éléments qui permettront de trancher sur la question des « hasards » du lexique en montrant que les syllabes ne répondent pas seulement à un pur formalisme, mais aussi à des contraintes de production et de perception de parole. Une telle démarche remet en cause les principes saussuriens d'arbitrarité des unités linguistiques, d'indépendance forme/substance et de primauté de la langue sur la parole dans les études linguistiques.

BIBLIOGRAPHIE

- BLEVINS, J. (1996) « The syllable in phonological theory » in *The Handbook of Phonological Theory*, Blackwell Publishers Ltd., Cambridge-Massachusetts, Oxford-UK : 206-244.
- BOË, L.J., VALLÉE, N., BADIN, P., SCHWARTZ, J.L. et ABRY, C. (2000) « Tendances in phonological structures : The influence of substance on form » in *Current Trends in Phonology and Phonetics II : Relationship between phonetics and phonology*, *Les Cahiers de l'ICP, Bulletin de la Communication Parlée*, 5 : 33-55.
- JANSON, T. (1986) « Cross-linguistic Trends in the Frequency of CV Sequences » in *Phonology Yearbook*, 3 : 179-195.
- KRAKOW, R. (1999) « Physiological organization of syllables : a review » in *J. of Phonetics*, 27 : 23-54.
- LINDBLOM, B. (2000) « Developmental origins of adult phonology : The interplay between phonetic emergents and the evolutionary adaptations of sound patterns » in *Phonetica*, 57/2-4 : 297-314.
- LINDBLOM, B., MACNEILAGE, P.F. et STUDDERT-KENNEDY, M. (1984) « Self-organizing processes and the explanation of languages universals » in *Explanation of Language Universals*, ed. Butterworth B., Comrie B. et Dahl O., Mouton : 181-203.
- LOCKE, J.L. (1983) *Phonological acquisition and change* ; Academic Press, New-York.
- MACNEILAGE, P.F. (1998) « The Frame/Content theory of evolution of speech production » in *Behavioral and Brain Sciences*, 21/4 : 499-511.
- MACNEILAGE, P.F. et DAVIS, B.L. (2000) « Deriving speech from non-speech : a view from ontogeny » in *Phonetica*, 57/2-4 : 284-296.
- MADDIESON, I. (1992) « The structure of segment sequences » in *UCLA Working Papers in Phonetics*, 83 : 1-8.
- MONINO, Y. et ROULON, P. (1972) « Phonologie du Gbaya Kara 'Bodoe » in *SELAF 31*, Paris.
- ROUSSET, I. (2000) *Des lexiques aux syllabes des langues du monde : données, typologie et tendances universelles* ; Mémoire de DEA en Sciences du Langage, Université Stendhal, Grenoble.
- SCHWARTZ, J.L., BOË, L.J., VALLÉE, N. et ABRY, C. (1997) « The Dispersion-Focalization Theory of Vowel Systems » in *J. of Phonetics*, 25 : 255-286.

TULLER, B. et KELSO, J.A.S. (1991) « The Production and Perception of Syllable Structure » *J. of Speech and Hearing Research*, 34/3 : 501-508.

VALLÉE, N. et BOË, L.J. (2000) « L'acquisition des structures sonores du langage par l'enfant et les tendances phonologiques des langues du monde » in *Revue de Linguistique et de Didactique des Langues Lidil*, 22 : *Acquisitions langagières et usages linguistiques enfantins* : 99-121.

VALLÉE, N., BOË, L.J. et STEFANUTO, M. (1999) « Typologies phonologiques et tendances universelles. Approche substantialiste » in *Linx, numéro spécial 1999* : 31-54.

VIHMAN, M.M. (1996) *Phonological development : The origin of language in the child* ; Blackwell Publishers Ltd, Oxford-UK.

AUTEURS

NATHALIE VALLÉE

Institut de la Communication Parlée, UMR 5009 INPG/Université Stendhal, BP 25 38040 Grenoble Cedex

Tél : 04 76 82 41 19 – Fax : 04 76 82 43 35 e-mail : vallee@icp.inpg.fr