

VICTOR ŽIVOV

UNE PROCÉDURE DE CLASSIFICATION DES CONSONNES DESTINÉE À LA DESCRIPTION DE LEURS COMBINAISONS

0. Dans la présente étude nous essayons de résoudre un problème particulier de la description phonologique et de montrer sur cet exemple l'une des approches possibles de ce problème qui se pose assez souvent quand il s'agit de formaliser la description linguistique.

Ce problème peut être résumé de la façon suivante: admettons que nous nous trouvons en présence d'un ensemble d'unités linguistiques, caractérisées par plusieurs indices. Pour décrire les modèles fonctionnels de ces unités, décrites préalablement par un ensemble d'indices donné, nous devons distribuer ces unités par classes non-intersectées et décrire ensuite leur fonctionnement non en termes d'unités isolées mais en termes de classes, auxquelles elles appartiennent. Les unités sont ordonnées selon le grade de similitude entre leurs caractéristiques, représentant les indices donnés. Alors, les problèmes suivants se posent: a) comment allons-nous évaluer le grade de similitude entre deux unités; b) comment définir le niveau critique de similitude, de façon à ce que deux unités données soient considérées comme similaires, si leur grade de similitude dépasse ce niveau, et comme non-similaires, s'il n'atteint pas ce niveau; c) puisque en général le rapport de similitude n'est pas transitif, nous devons organiser les classes de façon optimale par le rapport de tolérance. La réponse à toutes ces questions dépend naturellement de la nature des unités et des indices, donc on ne peut pas espérer construire une procédure tout à fait automatique, qui convienne à l'analyse de toutes les situations de ce genre. C'est pourquoi nous croyons qu'il serait souhaitable de décrire des procédures analogues pour différents systèmes linguistiques particuliers, en espérant que plus tard, quand un certain nombre de telles procédures sera élaboré, il sera possible de fixer leurs traits isomorphes et de formuler certaines restrictions linguistiques.

1. Nous nous occuperons du problème de la classification des combinaisons de consonnes russes (le russe étant une langue avec une syntagmatique de consonnes suffisamment complexe).¹

1.1. J. GREENBERG (1965) décrit différentes régularités typologiques concernant les combinaisons de consonnes sous forme d'implications, et démontra, en particulier, qu'un certain type de combinaison dans une langue peut en impliquer un autre. J. GREENBERG (1966 et 1969) et B. A. USPENSKY (1972) par ailleurs ont démontré que de telles régularités sur le plan typologique présupposaient des relations hiérarchiques à l'intérieur d'une même langue, ces dernières pouvant être beaucoup plus riches que les régularités typologiques (puisque toutes les relations hiérarchiques d'une langue n'ont pas leur équivalent typologique). Ceux qui parlent de la « difficulté de prononciation » de certaines combinaisons en l'expliquant par des raisons phonétiques générales s'appuient inconsciemment sur l'organisation hiérarchique de l'ensemble des combinaisons de consonnes dans la langue.² Cette explication générale est peu convaincante puisqu'elle ne peut pas être propagée à toutes les langues. Pourtant il s'agit ici, probablement, de différents types de combinaisons de consonnes et de l'ordre hiérarchique de ces types. Ces notions servent de base à la définition de « structural » et « accidental gaps ».

1.2. Quelle que soit la classification, *le type* de la combinaison de consonnes est toujours défini par *le type* de chacun de ces composants. Donc, le problème des types de combinaisons de consonnes peut être ramené au problème des types de consonnes relevant pour la description de leur syntagmatique (leur faculté de former des combinaisons avec d'autres consonnes). Ces types (classes) de consonnes doivent être organisés de telle manière que la distribution des consonnes appartenant au même type soit relativement semblable, et celle des consonnes appartenant aux types différents – relativement différente.

Une fois l'ordre des types de combinaisons de consonnes établi,

¹ Notons que les procédures de classification des consonnes proposées jusqu'à présent conviennent aux langues qui ont un système de combinaisons de consonnes relativement simple (v. J. KURYŁOWICZ, 1948; B. SIGURD, 1955; 1965), c'est pourquoi elles sont inefficaces pour le russe ou le géorgien, par exemple (v. H. VOGT, 1954, p. 32). Le choix de la langue joue donc un rôle très important.

² V. les différentes théories de ce genre, en commençant par la « théorie sonore » de O. JESPERSEN (1932), jusqu'aux constructions récentes de B. HÁLA (1956; 1961).

nous pouvons résoudre un grand nombre de problèmes linguistiques. C'est nécessaire avant tout pour la typologie, car cela permet de dire quels types de combinaisons sont caractéristiques pour la langue donnée. L'établissement d'une telle hiérarchie facilite aussi la délimitation des syllabes: on choisira parmi les décompositions possibles du groupe intervocalique, celle qui garantit aux deux parties obtenues la position la plus élevée dans l'ordre hiérarchique des combinaisons respectivement finales ou initiales.³ On pourra aussi résoudre le problème des combinaisons intervocaliques non-décomposables: ceci au cas où la combinaison ne peut pas être décomposée au niveau des combinaisons de consonnes concrètes, mais seulement au niveau de leurs types. Les dépendances hiérarchiques permettent dans certains cas d'expliquer des faits historiques (par exemple, l'assimilation de certaines combinaisons de consonnes et la réorganisation des autres dans le processus d'emprunt) et dialectologiques (v. V. M. ŽIVOV, 1971).

1.3. Quelle sorte de classification de consonnes conviendrait le mieux à la description de leurs combinaisons?

1.3.1. Les descriptions traditionnelles de la phonotactique se basent sur les classes phonétiques, qui comprennent le plus souvent les occlusives, les spirantes et les sonantes. On introduit aussi, si c'est nécessaire, des subdivisions plus étroites: le point d'articulation, la sonorité, l'aspiration, etc.; les sonantes se décomposent parfois en nasales et non-nasales, parfois aussi en demi-voyelles. Mais puisqu'il s'agit pour les langues différentes de classifications différentes, et que ces dernières ne correspondent en aucune manière aux particularités phonétiques de la langue donnée, nous pouvons en déduire que ces classifications ne doivent pas reposer sur une base phonétique. C'est ainsi qu'en arménien nous trouvons les sonantes en qualité de derniers éléments de la combinaison finale dans les groupes *-rm*, *-ym*, *-zm*, *-hm*, *-gm*, *-yr*, *-yn*, *-yl* (S. B. TOŠJAN, 1969, pp. 118-119); pour rendre la description de la distribution plus efficace, il aurait fallu, probablement, isoler /m/ de toutes les autres sonantes, /m/ et /n/ appartiendraient alors aux classes différentes.

³ La solution proposée par E. PULGRAM (1970) (v. encore W. LEHFELDT, 1971) nous semble trop artificielle, car ces auteurs font un choix mécanique entre les décompositions syllabiques possibles; « *the burden of irregularity* » d'après eux repose toujours sur le groupe final (*la coda*), ce qui a sa raison d'être pour les langues européennes, mais ne peut pas servir d'explication universelle, puisque pour certaines langues la structure des combinaisons finales est plus stricte que celle des combinaisons initiales.

Cette délimitation ne peut pas être expliquée par des raisons phonétiques, puisque le même critère (le point d'articulation) ne joue aucun rôle pour /p/ et /t/, par exemple, ayant des distributions très proches. D'autre part, pour beaucoup de langues la séparation de /m/ du /n/ n'a pas de sens.⁴

1.3.2. Donc, la description de la syntagmatique phonologique ne peut pas toujours se baser sur les données de la classification phonétique; nous sommes contraints de faire appel aux données-mêmes de la distribution lorsqu'il s'agit d'établir une classification de phonèmes, utile à la description de la syntagmatique phonologique. Il arrive souvent que ce problème ne soit pas posé explicitement, mais c'est de cela qu'il s'agit quand la question de la définition distributive du phonème est posée, car elle présuppose un inventaire de phonèmes fixé d'avance. Or, les méthodes de classification existantes ne sont point satisfaisantes: ou bien elles sont orientées vers les langues avec une syntagmatique de consonnes relativement simple, ou bien les résultats de leur application n'ont pas d'interprétation linguistique (v. H. VOGT, 1954).⁵

Donc, dans la présente étude nous nous apprêtons à proposer une classification de phonèmes, utile à la description de leur distribution et fondée sur leur distribution, et ceci en nous appuyant sur l'affinité distributive des phonèmes différents, cette dernière exprimant le grade de similitude des environnements de ces phonèmes. C'est pourquoi une classification distributive doit nécessairement reposer sur la comparaison des environnements des phonèmes et par conséquent sur l'évaluation de leur similitude; si le coefficient de similitude distributive de deux phonèmes donnés dépasse un certain niveau, ils entreront dans une même classe distributive, sinon – dans deux classes différentes. Si nous définissons les différentes classes de distribution comme les environnements de phonèmes différents nous obtiendrons pour les

⁴ Aussi, /m/ est opposé à toutes les autres sonantes dans d'autres langues indo-européennes, comme, par exemple, le norvégien (H. VOGT, 1942, pp. 16, 18), ou le serbo-croate (S. M. TOLSTAJA, 1972, pp. 6-8, 14). Le phonème /f/ s'oppose de la même manière par le caractère de sa distribution à /s/ et se rapproche des occlusives en margi (P. LADEFOGED, 1968, p. 65), et en norvégien (H. VOGT, 1942, p. 16). Les spirantes postlinguales se distinguent des spirantes prélinguales et se rapprochent des occlusives en géorgien (H. VOGT, 1958, p. 14) et pour beaucoup de langues turques.

⁵ V. F. W. HOUSEHOLDER (1962), où l'auteur propose des algorithmes pour donner des définitions des phonèmes anglais basées sur leur distribution; il mentionne ce fait lui-même. En même temps l'algorithme de Householder n'est autre chose que l'aboutissement logique de la méthode proposée par L. BLOOMFIELD (1933).

langues avec une syntagmatique complexe un ensemble de classes différentes. Mais les environnements de certains phonèmes sont plus ressemblants que d'autres, et par ces ressemblances et ces dissemblances ils caractérisent différemment les différents groupes de phonèmes, ce qui nous permet de fonder notre classification sur la similitude des environnements.

Nous trouvons la même approche dans les études de J. D. O'CONNOR et J. L. M. TRIM (1953) et de G. F. ARNOLD (1956; 1964), consacrées à la description des classes fonctionnelles de consonnes et de voyelles. La similitude distributive d'après ces auteurs correspond au nombre d'éléments communs dans les environnements des deux phonèmes (pris en rapport avec le nombre maximal d'éléments communs possible); les phonèmes sont considérés comme proches du point de vue de la distribution, si ce rapport dépasse un niveau arbitrairement fixé auparavant. Un tel procédé nous permet de séparer les consonnes des voyelles, mais il ne convient pas à la classification distributive des consonnes. Ce coefficient est efficace pour la comparaison de la distribution des consonnes et des voyelles, car elles ont des caractéristiques distributives très opposées. L'application de cette procédure aux combinaisons de consonnes (pour une langue slave, par exemple) ne donnera pas de résultats essentiellement différents pour différentes classes de consonnes, et, par conséquent, ne peut pas servir de base pour notre classification.

2. Mais pour les entourages nous pouvons fixer des évaluations moins approximatives; c'est ce que nous nous proposons de faire dans la présente étude, en basant là-dessus notre classification des consonnes. Nous allons opérer avec des groupes de deux consonnes initiales et finales. Nous croyons nécessaire d'avoir au départ des combinaisons formées par deux éléments et non celles qui comprennent plus de deux éléments comme le fait, par exemple, J. KURYŁOWICZ (1948; 1952), parce que toute procédure orientée vers les combinaisons plus longues n'est pas pour cette même raison universelle (v. E. FISCHER-JØRGENSEN, 1952, p. 25). Nous allons distinguer quatre positions: *I1* (position initiale suivie d'une voyelle); *I2* (position initiale suivie d'une consonne); *F1* (position finale, précédée d'une voyelle); *F2* (position finale précédée d'une consonne). Nous allons définir l'environnement d'un phonème donné dans une position donnée comme l'ensemble des consonnes, capables de former avec le phonème donné des composés binaires dans la position donnée. Le rapport de similitude distributive sera d'abord

défini pour chaque phonème pris dans toutes les positions possibles, puis sans l'évocation de la position.

2.1. Les phonèmes suivants pris dans la position *F2*, par exemple, ont les entourages suivants:

$$z': r, l'; \quad z: r, l, l'; \quad x: r, l', t, s.$$

Bien que le nombre de phonèmes communs dans tous ces groupes reste le même, il est évident que l'environnement du phonème z /*r, l, l*/ est plus proche de celui de z' /*r, l*/, que ne l'est celui de x /*r, l', t, s*/, ce qui est dû probablement au fait que le phonème /*l*/, faisant partie seulement de l'entourage de z , appartient au type phonétique que nous trouvons dans l'entourage de z' (les liquides), tandis que les phonèmes de l'entourage de x /*t, s*/ n'appartenant pas à l'entourage de z' , représentent un type phonétique nouveau (par rapport à l'entourage de z'). Ce point peut être précisé à l'aide de la méthode des traits distinctifs (TD): la différence entre le phonème /*l*/ et les phonèmes de l'entourage de z' se résume en un seul TD, tandis que le phonème /*s*/, qui fait partie de l'environnement de x , diffère de tous les phonèmes de l'entourage de z' au moins par trois TD. L'environnement de z' étant inclus dans les deux autres environnements, est aussi proche de l'un que de l'autre.⁶ Nous pouvons donc atteindre un grade de précision plus élevé en tenant compte de la variété phonétique des deux entourages donnés ou de leur valeur relative en termes de TD.

2.1.1. Formulons maintenant la définition du grade de similitude des phonèmes *A* et *B*, pris dans une position donnée.

Si $\overrightarrow{\Delta AB}$ est le grade de similitude de *A* et de *B*,

N_A est le nombre de phonèmes de l'entourage de *A*,

n_i est le nombre de phonèmes de l'entourage de *A* pour lesquels on pourra trouver dans l'entourage de *B* des phonèmes, qui diffèrent de ces derniers par *i* TD (pas moins que *i* TD), alors

$$\overrightarrow{\Delta AB} = \frac{\sum_{i=0}^8 \frac{n_i}{2^i}}{N_A}$$

⁶ Nous avons décrit l'inventaire des combinaisons de consonnes non-périphériques, ainsi que l'argumentation du système de notation adopté dans V. M. Živov (1973).

Nous pouvons nous servir de la formule approximative, sans tenir compte des termes quand $i > 2$, ce qui donnera

$$\overrightarrow{\Delta AB} = \frac{n_0 + \frac{n_1}{2} + \frac{n_2}{4}}{N_A}$$

Le grade de similitude de B par rapport à A sera défini de la même manière. Le grade de similitude de A et de B (ΔAB) est égal à

$$\Delta AB = \frac{\overrightarrow{\Delta AB} + \overleftarrow{\Delta AB}}{2}$$

Ce qui nous donne pour notre exemple:

$$\overrightarrow{\Delta z'z} = \overrightarrow{\Delta z'x} = \frac{2}{2} = 1 \quad \overleftarrow{\Delta z'x} = \frac{2 + \frac{1}{2}}{4} = 0,625$$

$$\overleftarrow{\Delta z'z} = \frac{2 + \frac{1}{2}}{3} = 0,833 \quad \Delta z'z = \frac{1}{2} \cdot (1 + 0,833) = 0,92$$

$$\Delta z'x = \frac{1}{2} \cdot (1 + 0,625) = 0,81$$

Ni la première ni la seconde formule ne sont symétriques, ce qui répond au fait que l'entourage de z' est également proche de ceux de z et de x , mais les entourages de x et de z se trouvent à distances différentes de celui de z' .

2.2. C'est ainsi que pour chaque phonème nous obtiendrons trois coefficients: le grade de similitude de A par rapport à B , de B par rapport à A et de A et B . Grâce à ces indices nous pouvons obtenir des coefficients relatifs; par exemple, pour dire que dans la position $F1$ $/s/$ est plus proche de $/r/$ que de $/l/$, etc. En partant de là il faut que nous passions à la définition du rapport absolu, pour pouvoir dire que dans une position donnée A et B sont proches, mais A et C ne le sont pas. Pour cela, d'autre part, il faut décider dans quels cas et de quels coefficients nous allons tenir compte.

Pour les phonèmes suivants dans *F2*, par exemple, nous avons les entourages suivants:

s: *k*
m: *j, d', n, s, š, r, l, l'*
v: *t, s, z, č, š, k, r, l*

L'entourage de */m/* n'a que quatre phonèmes en commun avec celui de */v/*, tandis que l'entourage de */s/* est inclu dans celui de */v/*. Pouvons-nous dire que */s/* est proche de */v/* de la même façon que */m/* est proche de */v/*? Il est souhaitable de parler de la similitude de *A* et de *B* quand les entourages présentent un volume à peu près égal; lorsqu'il s'agit d'entourages de volumes très différents nous parlerons du grade d'emboîtement de l'entourage de *A* dans l'entourage de *B*, c'est-à-dire de la similitude de *A* par rapport à *B*.⁷

Il serait trop facile de comprendre la corrélation entre les volumes comme une proportion entre le nombre de phonèmes, formant les entourages. Nous allons définir le volume comme le nombre de types phonétiques représentés dans l'environnement de phonème donné, c'est-à-dire la variété phonétique de l'entourage.

C'est ainsi que dans la position *F1*, par exemple, l'entourage de *b* sera *r, r', l, l'*, celui de *v*: *d, n, r, l*. L'environnement de *b* est composé uniquement de liquides, tandis que celui de *v* est plus varié. Pour choisir entre ces deux entourages (de *A* et de *B*) celui qui présente plus de variété phonétique nous pouvons utiliser la formule évaluant le grade de la similitude.

2.2.1. Maintenant nous allons définir un nouveau rapport:

x est plus riche que γ (inclu $\gamma, \gamma \vdash x$) $\equiv \Delta x \overrightarrow{\gamma} < \overleftarrow{\Delta x \gamma}$.

Si nous rangeons les consonnes dans chacune des positions relativement à ce rapport, nous obtiendrons les résultats suivants (les échelles de l'inclusion):

⁷ Ce problème gagne d'importance quand il s'agit de la distribution défectueuse. Dans la position *F2* les entourages de *b, b', t* et *t'* sont respectivement *b*: *d', z, ž, r, l*; *b'*: *r*; *t*: *j, s, p, č, r, l*, et *t'*: *s, r, l*. L'entourage du phonème marqué est une sorte d'entourage réduit du phonème non marqué. Si nous évaluons ces paires de phonèmes comme nous le ferions pour la paire $\langle r, l' \rangle$, par exemple, dans la position *F1* (*r*: *t, d, n, s, z, p, b, m, v, ž, k, g, x*; *l'*: *t, d, s, z, p, d, m, v, č, š, k, g, x*) ces phonèmes seront répartis dans des classes différentes (elles seront donc non-similaires).

I1: $p, p', z, z', x, b' \vdash b \vdash s', s, \xi: \vdash k \vdash \xi, \xi \vdash d \vdash \xi \vdash g \vdash t' \vdash d' \vdash t, m \vdash m', j \vdash v' \vdash n' \vdash l \vdash r \vdash n \vdash v \vdash r', l'.$

I2: $c, n, t' \vdash l \vdash \xi \vdash g \vdash f \vdash b \vdash x \vdash r \vdash k \vdash d \vdash z \vdash l' \vdash t \vdash v \vdash \xi \vdash \xi \vdash m \vdash p, s.$

F1: $f, x, r' \vdash n, d' \vdash j \vdash m \vdash d \vdash \xi \vdash b \vdash \xi \vdash v \vdash g \vdash p \vdash k \vdash \xi \vdash l', t, z \vdash s \vdash l \vdash r.$

F2: $c, s, b', \xi \vdash z' \vdash z \vdash \xi \vdash \xi: \vdash p, t', k \vdash r' \vdash n' \vdash g \vdash v' \vdash \xi \vdash l' \vdash x \vdash d' \vdash b \vdash d \vdash r \vdash t \vdash n \vdash m \vdash v \vdash l.$

Pour la langue russe le rapport $x \vdash y$ n'est pas transitif, car dans certains cas $x \vdash y, y \vdash z, z \vdash x$ (les entourages ne se superposent qu'en partie). Ces phonèmes seront considérés comme membres d'un même niveau, ou bien le plus petit des coefficients sera éliminé, puisque dans ce cas les entourages ne se confrontent qu'en partie (la procédure exclut les paires de phonèmes qui se diffèrent par plus de deux traits distinctifs). De même, des phonèmes x et y seront considérés comme les éléments de même niveau, si $\Delta \overrightarrow{xy} = \Delta \overleftarrow{xy}$.

2.2.2. A présent nous déciderons dans quels cas les volumes seront considérés suffisamment proches, pour que la comparaison s'effectue à l'aide du coefficient symétrique (ΔAB) , et dans quels cas nous les tiendrons pour suffisamment dissemblants pour être mesurés avec le coefficient orienté $(\Delta \overrightarrow{AB})$. Puisque nous ne voyons pas de critères linguistiques valables, nous allons arbitrairement partager l'échelle, par exemple en trois parties; nous pouvons attribuer à chaque rang la valeur D .

Soit p - le nombre de phonèmes dans l'échelle
 q - le nombre de phonèmes, formant le niveau donné
 r - le nombre de phonèmes, formant les niveaux inférieurs au niveau donné.

On attribuera alors au niveau donné (aux phonèmes du niveau donné) l'évaluation

$$D = \sum_{i=1}^q \frac{r+i}{pq}$$

*

Les phonèmes de chaque rang peuvent être séparés en phonèmes-contenants ($D \geq 2/3$), moyens ($2/3 > D \geq 1/3$) et inclus ($D > 1/3$). Nous allons noter ces groupes 2, 1, 0.

Nous allons définir maintenant pour chaque position les moyennes suivantes:

soit p_2 - le nombre des phonèmes-contenants
 p_1 - le nombre des phonèmes moyens
 p_0 - le nombre des phonèmes inclus

$$\begin{aligned} \overline{\Delta x_2 \gamma_2} &= \frac{2 \sum \Delta x_2 \gamma_2}{p_2(p_2 - 1)} & \overline{\Delta x_1 \gamma_2} &= \frac{\sum \overrightarrow{\Delta x_1 \gamma_2}}{p_1 p_2} \\ \overline{\Delta x_1 \gamma_1} &= \frac{2 \sum \Delta x_1 \gamma_1}{p_1(p_1 - 1)} & \overline{\Delta x_0 \gamma_2} &= \frac{\sum \overrightarrow{\Delta x_0 \gamma_2}}{p_0 p_2} \\ \overline{\Delta x_0 \gamma_0} &= \frac{2 \sum \Delta x_0 \gamma_0}{p_0(p_0 - 1)} & \overline{\Delta x_0 \gamma_1} &= \frac{\sum \overrightarrow{\Delta x_0 \gamma_1}}{p_0 p_1} \end{aligned}$$

Soit A et B deux phonèmes-contenants, moyens ou inclus. A et B sont proches, si

$$(1) \quad \Delta A_i B_i > \Delta x_i \gamma_i$$

Soit deux phonèmes A et B qui appartiennent à deux sections différentes; A appartient au groupe, situé plus bas que le groupe, auquel appartient B (B contient A). A est proche de B , si

$$(2) \quad \Delta \overrightarrow{A_i B_u} > \overline{\Delta x_i \gamma_u}$$

2.3. Nous dirons que A et B sont proches, s'ils sont proches (v. (1) et (2)) dans toutes les positions.

On obtient de cette façon un ensemble de groupes fermés par le rapport de similitude, c'est-à-dire des groupes dont tout élément est proche de tous les autres éléments du même groupe. Puisque le rapport de similitude tout en n'étant pas transitif est toutefois tolérant un même phonème peut appartenir à plusieurs groupes différents. Voici l'ensemble de groupes fermés, obtenu pour la langue russe:⁸

⁸ Les phonèmes p' , b' , c , s' , z' ne figureront pas dans cette classification, car leur distribution, qui est défectueuse, ne les caractérise en aucune façon (p' n'entre que dans une

- | | |
|-----------------------|--------------------------------|
| 1. p, d, d', b | 15. z, d, j |
| 2. p, d, d', v | 16. $z, \text{ž}, j$ |
| 3. p, d', t', v | 17. g, b, b' |
| 4. b, g, d, d' | 18. $\text{š}', g, x$ |
| 5. t, t', d', k | 19. $\text{š}', g, b'$ |
| 6. t, d, d' | 20. p, m, m' |
| 7. $t', d', \text{ž}$ | 21. p, v, v' |
| 8. d, d', n | 22. p, v, m' |
| 9. n, n', d | 23. p, f |
| 10. n', d, g | 24. m', k |
| 11. n', d, p | 25. v', k |
| 12. n', g, x | 26. $s, z, \text{š}, \text{ž}$ |
| 13. $d', j, \text{ž}$ | 27. $s, z, \text{ž}$ |
| 14. d', d, j | 28. r, r', l, l' |

2.4. Maintenant, après ces groupes nous devons passer à la division en classes, de telle façon que chaque phonème n'appartienne qu'à une seule classe. L'étape suivante sera donc la procédure de regroupement des groupes fermés et peut se baser sur les facteurs suivants: (a) quelle est la moyenne qui caractérise l'affinité de la distribution positionnelle (du point de vue de la distribution des éléments dans les classes des tenants, des phonèmes moyens et des phonèmes inclus, pris dans les 4 positions); les groupes avec une moyenne d'affinité mineure peuvent être ignorés; (b) à quel point les éléments dans le groupe sont-ils liés entre eux, c'est-à-dire quelle est la moyenne de similitude de tous les éléments du groupe; les groupes, caractérisés par une moyenne de similitude basse peuvent être ignorés; (c) quels changements provoque l'adjonction d'un élément donné ou son exclusion dans les paramètres (a) et (b); si la moyenne de similitude ou la moyenne d'affinité positionnelle se trouvent très diminuées par ce fait, l'opération sera considérée comme non applicable, sinon elle est à appliquer; (d) il est naturel de vouloir arranger les éléments de telle manière que les classes soient suffisamment vastes et d'autre part qu'il n'y ait pas de classes à un seul élé-

seule combinaison $sp'-$; b' - dans deux - lb' - et $-rb'$; c dans cv' - et $-šc$; s' dans ps' - et $fs'-$; z' dans $vz'-$). Les données de la distribution permettent de faire inclure ces phonèmes dans des classes différentes. Donc, en accord avec la distribution, nous pouvons les placer avec les phonèmes qui leur sont phonétiquement les plus proches: p avec p' , b avec b' , s' avec s , z' avec z . Pour c la question reste ouverte ainsi que pour f et $\text{š}'$, comme nous le verrons par la suite.

ment, qui de son côté soit uni à un autre phonème par le rapport de similitude maximale.

Nous pouvons proposer deux procédures basées respectivement sur des paramètres de type (a) et de type (b). Leurs résultats sont équivalents (à l'exception d'un détail peu important).

2.4.1. Nous pouvons faire correspondre à chaque phonème un cortège de quatre éléments $\langle a_1, a_2, a_3, a_4 \rangle$ où a_1 est le groupe d'éléments auquel appartient le phonème donné dans la position $I1$ (phonèmes-contenants - 2, moyens - 1, ou inclus - 0), a_2 - la même chose pour $I2$, a_3 pour $F1$ et a_4 pour $F2$. Nous allons définir la différence entre les distributions positionnelles (δ) de deux phonèmes, auxquels correspondent les cortèges $\langle a_1, a_2, a_3, a_4 \rangle$ et $\langle b_1, b_2, b_3, b_4 \rangle$ comme:

$$\delta_{ab} = \sum_{i=1}^4 |a_i - b_i|.$$

Toujours $0 \leq \delta \leq 8$.

En définissant la différence entre les distributions positionnelles il faut tenir compte des particularités des phonèmes défectueux. Les consonnes palatales russes (excepté l' , ξ' et en partie ξ inclus) ne peuvent être suivies dans les groupes tautosyllabiques que de consonnes labiales; c'est pourquoi elles font partie de la catégorie des phonèmes inclus dans $I2$ et $F1$, ce qui les caractérise en tant qu'un tout mais pas sur le plan individuel. En vue de cet état de choses nous introduisons deux coefficients: le premier (δ_1) évalue la moyenne de différence positionnelle seulement entre les phonèmes non-défectueux, le second (δ_2) pour tous les phonèmes quels qu'ils soient. Si le groupe se compose de n phonèmes, dont m ne sont pas défectueux, en notant les phonèmes du groupe $x_1 \dots x_m \dots x_n$, nous pouvons formuler les définitions suivantes:

$$\delta_1 = \frac{\delta_{x_1 x_2} + \dots + \delta_{x_1 x_m} + \delta_{x_2 x_3} + \dots + \delta_{x_2 x_m} + \dots + \delta_{x_{m-1} x_m}}{\frac{1}{2} m (m - 1)}$$

$$\delta_2 = \frac{\delta_{x_1 x_2} + \dots + \delta_{x_1 x_n} + \delta_{x_2 x_3} + \dots + \delta_{x_2 x_n} + \dots + \delta_{x_{n-1} x_n}}{\frac{1}{2} n (n - 1)}$$

Pour certains groupes le coefficient δ_1 ne peut pas être calculé (quand $m \leq 1$).

A présent nous pouvons en utilisant δ_1 et δ_2 définir pour tous les groupes la moyenne de différence positionnelle à l'intérieur du groupe; nous obtiendrons les coefficients $\bar{\delta}_1$ et $\bar{\delta}_2$. Après cela nous pouvons éliminer tous les groupes pour lesquels $\delta_1 > \bar{\delta}_1$ & $\delta_2 > \bar{\delta}_2$ ou au cas où il n'y aurait pas $\delta_1, \delta_2 > \bar{\delta}_2$. Donc nous pouvons rayer de la liste les groupes 3, 11, 13, 15, 16, 20, 21, 22, 23, 24, et 25.

Maintenant réunissons les groupes pour lesquels $\delta_1 < \bar{\delta}_1$ et $\delta_2 < \bar{\delta}_2$ qui contiennent le même ensemble d'éléments non-défectueux. Cette opération est acceptable si pour les groupes obtenus $\delta_1 < \bar{\delta}_1$ & $\delta_2 < \bar{\delta}_2$. De cette façon nous pouvons fusionner les groupes 1, 4, 5, 6, 10, et d'autre part 26 et 27. Si l'opération n'est pas applicable il faut éliminer les éléments qui nous empêchent de l'effectuer (nous ne discutons pas les difficultés qui peuvent surgir ici car elles ne se sont pas présentées pour la langue russe). Maintenant occupons-nous des groupes pour lesquels $\delta_1 < \bar{\delta}_1 \vee \delta_2 < \bar{\delta}_2$. Voyons s'il est possible de faire adjoindre aux groupes nouvellement formés qui contiennent les mêmes éléments non-défectueux, ceux de leurs éléments qui ne sont encore rattachés à aucune classe. On essaiera d'abord de rattacher les phonèmes non-défectueux puis les phonèmes defectueux. Si après avoir inclus un phonème non-défectueux nous constatons que pour le groupe formé de cette manière $\delta_1 > \bar{\delta}_1$ ou $\delta_2 > \bar{\delta}_2$, l'adjonction n'aura pas lieu. On choisira dans l'ensemble des phonèmes pour lesquels l'adjonction est possible celui dont l'inclusion donne au coefficient $\bar{\delta}_1 - \delta_1$ la valeur minimale. Cette classe nouvellement formée est soustraite à son tour à la même procédure, appliquée aux phonèmes, qu'on pouvait inclure dans la classe précédente. Conformément au facteur (d) (v. 2.4.) voyons si parmi les phonèmes defectueux qui appartiennent à la classe obtenue ou qui pouvaient y être inclus il s'en trouve un qui soit lié par le rapport de similitude (Δxy) maximale à un phonème non-défectueux, l'inclusion duquel dans la classe donnée se présenterait comme impossible. Maintenant on analysera les possibilités d'inclusion pour les phonèmes restants au moyen de la même procédure (à la seule différence, que l'on n'utilisera plus que le coefficient δ_2).

Après avoir appliqué cette procédure nous obtiendrons les classes suivantes: $\{p, d, d', b, g, t, t', k\}$, $\{n, n'\}$, $\{s, z, š, ž, č, \}$, $\{r, r', l, l'\}$.

Il nous reste à classer les phonèmes suivants: $v, v', m, m', x, j, š', f$. Le cas de p', b', s', z', c a été discuté plus haut. Dans le cadre de la procédure donnée il nous est impossible de classer $x, j, š', f$ (v. à ce sujet 3.1.). Quant à v, v', m, m' , nous pouvons obtenir les groupes $\{v, v'\}$, $\{v, m'\}$, $\{v, v'\}$ pour lesquels $\delta_2 < \bar{\delta}_2$, après avoir éliminé $\{p/$

des groupes 20, 21 et 22. Puis en conformité avec le rapport de similitude nous choisirons $\{m, m'\}$ et $\{v, v'\}$, parce que ces éléments sont liés plus étroitement entre eux. Après cela nous pouvons considérer l'application de la procédure comme terminée.

2.4.2. La deuxième procédure possible est basée sur le facteur de la moyenne de similitude à l'intérieur du groupe; il faut donc définir le coefficient de similitude absolue pour deux phonèmes, c'est-à-dire indépendamment de leur position; il serait naturel de la définir comme la moyenne des coefficients symétriques de similitude pour les quatre positions possibles. Ceci est vrai, mais avec certaines réserves. Si dans une position donnée les deux phonèmes appartiennent à la catégorie des inclus, leur environnement ne les caractérise en aucune manière, donc il y a des raisons pour ignorer le coefficient qui correspond à cette position.⁹ Ou encore si le phonème est défectueux, il aura en cette position un environnement nul ou bien un environnement qui ne le caractérisera en aucune manière; c'est pourquoi ces positions défectueuses doivent être omises quand il s'agit de définir la dite moyenne. Si deux phonèmes non-défectueux en une position donnée font partie l'un de la catégorie des éléments-contenants, l'autre de celle des inclus, le coefficient symétrique est peu utilisable pour caractériser leur similitude. D'autre part la différence entre les volumes des entourages les caractérise comme non-similaires. C'est pourquoi on peut dire que le coefficient de similitude en ce cas est égal à zéro.

Compte tenu de ces objections on peut définir pour chaque paire de phonèmes le coefficient de similitude absolu, indépendamment de leur position ($\Delta'xy$) comme la somme des coefficients de similitude symétrique (Δxy) pris dans toutes les positions, divisée par le nombre de positions qui entrent en jeu.

Maintenant on pourra définir la moyenne de similitude pour les phonèmes d'un groupe fermé (nous avons procédé de la même manière avec le coefficient de différence positionnelle); après cela on fixera la moyenne des coefficients de similitude qui caractérisent tous les groupes fermés, après quoi les groupes avec $\Delta' < \bar{\Delta}'$ seront éliminés (dans notre exemple ce sera le cas de 7, 11, 12, 13, 15, 16, 19, 20, 21, 22, 23).

⁹ Pour les points obscurs on pourra faire appel directement aux échelles d'inclusions et analyser le coefficient de similitude aussi au cas où les phonèmes feraient partie de niveaux suffisamment proches, qui en même temps ne seraient pas les niveaux inférieurs de l'échelle.

On peut réunir au premier stade les groupes avec les mêmes éléments de façon différente. Pour commencer on peut unir deux groupes avec les coefficients les plus élevés (Δ'), et voir si pour eux $\Delta' > \bar{\Delta}'$. Nous pouvons proposer quelques autres variantes. Les phonèmes des groupes qui restent (de 1 à 18) seront distribués dans la même classe si chacun d'eux fait partie de plus de deux groupes. Pour le groupe ainsi formé nous aurons $\Delta' > \bar{\Delta}'$. L'algorithme d'adjonction pour les autres phonèmes est le même que dans la première procédure. En tenant compte du facteur (d) (v. 2.4.) et des groupes 20-21, pour lesquels après l'élimination de /p/ $\Delta' > \bar{\Delta}'$, nous obtiendrons en fin de compte la classification suivante:

- I: *r, r', l, l'*
- II: *n, n'*
- III: *s, z, š, ž, č; s', z'*
- IV: *v, v'*
- V: *p, t, k, b, d, g, x, t', d'; p', b'*
- VI: *m, m'*

Nous avons placé les phonèmes défectueux après le signe « ; », ils sont rattachés aux classes correspondantes grâce à leur traits phonétiques quand leur distribution ne s'y oppose pas. Nous n'avons pas réussi à classer les phonèmes *j, š', f*, et *c*. La différence entre les résultats des deux procédures comme on voit ne concerne que le phonème /x/, que nous n'avons pas réussi à classer dans le cadre de la première procédure.

2.5. Maintenant nous pouvons appliquer aux classes obtenues le critère du nombre mineur de combinaisons internes (1) et celui de l'absence de paires symétriques (2).

Le premier critère peut être formalisé de la façon suivante:

a) Le nombre de combinaisons internes dans le groupe doit être plus petit que le nombre d'éléments qui appartiennent à ce groupe.

b) La valeur de la proportion entre le nombre de combinaisons internes et le nombre de combinaisons internes théoriquement possibles doit être plus petite que celle de la proportion entre le nombre de combinaisons de la classe donnée avec toute autre classe et le nombre de combinaisons de cette espèce théoriquement possible.

c) La classe ne doit pas contenir d'éléments qui après être séparés de leur classe et inclus dans une classe à part se combindraient pour la plupart avec les éléments de leur classe de départ.

Les classes obtenues ne violent aucune des conditions mentionnées plus haut; elles ne contiennent pas non plus de paires symétriques.

3.1. Comme nous l'avons vu, la présente procédure permet de classer les consonnes qui participent à un grand nombre de combinaisons, tandis que celles qui se combinent peu restent hors de son action. Ce fait n'est pas dû au hasard.

Abstrayons-nous de la présente procédure et essayons de nous représenter l'organisation des classes destinées à la description de la syntagmatique des consonnes. Chaque classe distributive se combine avec d'autres classes de phonèmes. Les différentes espèces de ces combinaisons s'organisent par rapport à chaque classe de façon hiérarchique.

Les phonèmes qui appartiennent à une classe donnée doivent se combiner avec les types hiérarchiques supérieurs de façon homogène; ils peuvent former avec les types inférieurs des combinaisons différentes. D'autre part les phonèmes de classes différentes se combinent différemment avec les types supérieurs; leurs combinaisons avec les types inférieurs peuvent en principe coïncider. Une telle classification hiérarchique permet de distinguer les traits caractéristiques dans la distribution du phonème donné et de classer de cette façon les phonèmes qui entrent dans peu de combinaisons.¹⁰ Mais la classification fonctionnelle présuppose la décomposition en classes distributives. C'est pourquoi la présente classification, basée sur les affinités phonétiques, doit précéder la classification fonctionnelle, qui complète la première et permet de classer certains phonèmes qu'on n'a pas pu classer au premier stade, ainsi qu'unir certaines classes (en utilisant la coïncidence des types hiérarchiques supérieurs).

3.2. Maintenant nous pouvons formuler certaines remarques générales concernant la procédure proposée ci-dessus.

Une telle procédure s'avère indispensable quand l'établissement de la hiérarchie fonctionnelle des éléments doit être précédé par leur classification préliminaire. Le plus souvent ces problèmes se posent dans le domaine de la phonologie et de la sémantique; quant aux problèmes morphologiques et syntaxiques, cette classification résulte de la corrélation du plan de l'expression et celui du contenu (par exemple, la hié-

¹⁰ Le phonème /x/ attaché à la classe V dans le cadre de la deuxième procédure, la première n'ayant pas pu le classer se retrouvera dans cette même classe aussi dans ce dernier cas après la classification fonctionnelle.

rarchie des fonctions d'un cas donné, ou la hiérarchie des cas, assumant cette fonction). Les classifications extérieures comme la nôtre ont un caractère préliminaire; elles peuvent subir des changements par la suite au cours de la classification fonctionnelle. Toute la procédure est donc organisée si l'on peut dire de façon cyclique (v. H. VOGT, 1942, pp. 6-7).

Les procédures comme la nôtre rendent possible la formalisation de tels concepts comme « le type de combinaison » pour les éléments phonétiques aussi bien que pour les éléments sémantiques; elles sont avec ça faciles à réécrire sous forme d'algorithme et peuvent en principe être appliquées par les machines électroniques. Pour notre procédure il aurait été suffisant d'introduire les données suivantes: la liste des combinaisons de consonnes, la matrice d'identification des consonnes en termes de traits distinctifs, les données, concernant les traits qui causent la déféctuosité. Il est peu probable que ces procédures donneront des résultats intéressants au point de vue linguistique au cas où elles seraient conçues sans tenir compte des traits spécifiques des unités linguistiques données (pour nous ce fut, par exemple, la déféctuosité de la distribution, etc.); mais on peut certifier que certains mécanismes particuliers comme l'algorithme d'adjonction de l'élément nouveau, seront universels pour toutes les procédures de ce genre.

BIBLIOGRAPHIE

- G. F. ARNOLD, *A Phonological Approach to Vowel, Consonant, and Syllable in Modern French*, dans «Lingua», V (1955-56), pp. 253-287.
- G. F. ARNOLD, *Vowel and Consonant - a Phonological Definition Re-examined*, dans *In Honour of Daniel Jones*, London, 1964.
- L. BLOOMFIELD, *Language*, New York, 1933.
- E. FISCHER-JØRGENSEN, *On the Definition of Phoneme Categories on a Distributional Basis*, dans «Acta Linguistica», VII (1952), pp. 8-39.
- J. H. GREENBERG, *Some Generalizations Concerning Initial and Final Consonant Sequences*, dans «Linguistics», XVIII (nov. 1965), pp. 5-34.
- J. H. GREENBERG, *Language Universals - with Special Reference to Feature Hierarchies*, The Hague, 1966.
- J. H. GREENBERG, *Methods of Dynamic Comparison in Linguistics*, dans *Substance and Structure of Language*, Los Angeles and Berkeley, 1969.
- B. HÁLA, *Slabica, její podstata a vyvoj*, Praha, 1956.
- B. HÁLA, *La syllabe, sa nature, son origine et ses transformations*, dans «Orbis», X (1961), pp. 69-143.
- F. W. HOUSEHOLDER, *The Distributional Determination of English Phonemes*, dans «Lingua», II (1962), pp. 186-191.
- O. JESPERSEN, *Lehrbuch der Phonetik*, 1932 (1^{er} edit. 1904).
- J. KURYŁOWICZ, *Contribution à la théorie de la syllabe*, 1948, dans *Esquisses linguistiques*, Wrocław-Kraków, 1960, pp. 193-220.
- J. KURYŁOWICZ, *Uwagi o polskich grupach spółgłoskowych*, 1952, dans *Esquisses linguistiques*, Wrocław-Kraków, 1960, pp. 221-232.
- P. LADEFOGED, *A Phonetic Study of West African Languages*, 1968^a.
- W. LEHFELDT, *Ein Algorithmus zur automatischen Silbentrennung*, dans «Phonetica», XXIV (1971), pp. 212-237.
- J. D. O'CONNOR, J. L. M. TRIM, *Vowel, Consonant and Syllable - a Phonological Definition*, dans «Word», IX (1953), pp. 103-122.
- E. PULGRAM, *Syllable, word, nexus, cursus*, The Hague, 1970.
- B. SIGURD, *Rank Order of Consonant Established by Distributional Criteria*, dans «Studia Linguistica», IX (1955), pp. 8-20.
- B. SIGURD, *Phonotactic Structures in Swedish*, Lund, 1965.
- S. M. TOLSTAJA, *Načal'nye i konečnye sočtanija soglasnych v serbohorvatskom jazyke. Issledovanija po serbohorvatskomu jazyku*, Moskva, 1972.
- S. B. TOŠJAN, *Slogorazdel i strojenje sloga v armjanskom jazyke*, dans «Voprosy Jazykoznanija», (1969) 3, pp. 116-124.
- B. A. USPENSKY, *Subsystems in Language, their Interrelations and their Correlated Universals*, dans «Linguistics», LXXXVIII (aug. 15, 1972), pp. 53-71.
- H. VOGT, *The Structure of the Norwegian Monosyllables*, dans «Norsk Tidsskrift

- for Sprogvidenskap », XII (1942), pp. 5-29.
- H. VOGT, *Phoneme Classes and Phoneme Classification*, dans « Word », X (1954), pp. 28-34.
- H. VOGT, *Structure phonémique du Géorgien*, dans « Norsk Tidsskrift for Sprogvidenskap », XVIII (1958), pp. 5-90.
- V. M. ŽIVOV, *Tipologija sočtanij soglasnych v russkich govorach*, dans « Voprosy Jazykoznanija », (1971) 2, pp. 69-82.
- V. M. ŽIVOV, *Centr i periferija v fonologičeskoj organizacii slova*, dans *Lingvotipologičeskie issledovanija*, part. I, Moskva, 1973.

