

LA VITESSE D'ARTICULATION ET LES UNITÉS SONORES
DANS LA CHAÎNE PARLÉE

M. Dohalská-Zichová

Institut de Phonétique, Université Charles, Prague.

ABSTRACT

The carefully arranged material represents a set of spontaneous announcements in natural working communication uttered continuously and with ever increasing speed until the articulation maximum is reached. It is used to examine the phenomena affecting intelligibility. A preliminary aural analysis is followed by a spectrographic study with special focus on the problem of linear and non-linear shortening of sounds, groups and possibly the influence of stress. The results of our research will be used to further improve the synthetic speech signal in Czech.

Les résultats de nos recherches précédentes concernant l'intelligibilité de la parole spontanée sur différents lieux de travail ainsi que la perfection de la qualité de la parole synthétique du tchèque ont inspiré la réalisation de l'expérience suivante.

Le point de départ de cette recherche représente un choix de 50 phrases tirées de la communication spontanée des réseaux de communication sur les différents lieux de travail. Ces 50 phrases ont été enregistrées par 8 speakers à qui on a demandé d'articuler assez distinctement pour que les phrases soient compréhensibles dans un débit accéléré, mais de garder en même temps la prononciation naturelle d'une communication quotidienne cou-

rante. Chaque phrase a été prononcée d'abord dans un débit "ordinaire" et, après une petite pause, le locuteur a répété plusieurs fois la phrase en augmentant continuellement le débit de l'articulation jusqu'au point où il ne peut plus répéter nettement la phrase. Ce type de réalisation sonore, tout à fait naturelle et différente de chaque accélération "mécanique", nous a permis de nous poser plusieurs questions concernant, d'une part, les problèmes de perception et, de l'autre, les changements dans l'image spectrographique.

Dans la première phase de nos investigations, nous nous sommes concentrés d'abord à l'analyse auditive (collective et individuelle). Il nous intéressait de savoir quel type de changements les auditeurs perçoivent au cours de l'augmentation de la vitesse de la parole.

Sur la base de sondages préliminaires, nous avons fait un enregistrement spectrographique représentatif (le matériel sonore dure plus de 80 min., les enregistrements spectrographiques représentent plus de 10 000 données). Ici, je voudrais remercier, M. M. Ptáček et Mlle P. Žáčková pour leur fructueuse collaboration. Nous avons voulu confronter d'abord les matériaux spectrographiques avec les acquis tirés de l'analyse auditive et nous nous sommes posés encore d'autres questions, comme p. e. :
- quelle est l'influence de l'accent sur la durée des voyelles?

- est-ce qu'il y a un parallèle dans l'abréviation des voyelles brèves et longues? (La longueur des voyelles en tchèque est fonctionnelle!)

- comment se reflète l'augmentation de la vitesse du débit dans les petits groupes rythmiques de la phrase?

- en quoi la forme sonore de la phrase prononcée le plus lentement est différente de celle qui est prononcée le plus vite?

Comme nous connaissons les travaux remarquables de Klatt, Umeda, Huggins et d'autres, comme les analyses fondamentales ont été réalisées sur le matériel du tchèque soutenu par M. Maláč et ses collaborateurs et comme, nous-mêmes, nous nous occupons de l'analyse spectrographique liée à la perception depuis longtemps, nous avons espéré que les sonogrammes seraient analysables facilement aussi bien par les méthodes statistiques ce qui nous permettrait de confronter, non seulement, les résultats de l'analyse auditive et spectrographique, mais d'obtenir aussi de nouveaux résultats contribuant au perfectionnement de la qualité de la synthèse de la parole tchèque.

Au cours de notre travail, nous avons vu qu'on ne pouvait y appliquer les critères utilisés pour l'élaboration des matériaux "de laboratoire", car il s'agit plus ou moins de la parole naturelle ainsi que de l'accélération naturelle, ce qui pose de problèmes nouveaux.

S'il faut considérer les paramètres influençant la durée des sons, nous devons constater que même quand il s'agit du même locuteur, la voyelle ou consonne donnée ne se comporte pas de la même façon dans le même entourage, dans le même contexte, si on augmente le débit de la parole. Chez les différents locuteurs, on peut constater des changements similaires, différents, mais aussi contradictoires - et ceci dans

les circonstances adéquates. P.e. /a/ accentué dans la syllabe initiale se montre ainsi (5 locuteurs, 3 réalisations accélérées continuellement; durée en "ms"):

55	40	40	45	65
55	40	45	40	45
40	40	45	40	40

Dans la syllabe accentuée au milieu de la phrase les changements dus au débit sont les suivants:

45	50	40	40	50
45	40	35	30	60
45	40	30	55	40

La relation du /i/ accentué et du /i/ inaccentué dans les syllabes successives se montre assez intéressante: dans la première réalisation non-accelérée le /i/ accentué est plus long ou semblable au /i/ inaccentué: 60:45, 35:35, 30:30, 45:40, 40:40

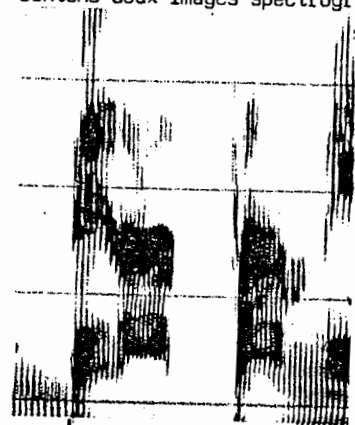
- dans les deux réalisations accélérées suivantes, se manifeste, au contraire, l'abréviation du /i/ accentué et l'allongement du /i/ inaccentué; la relation du /i/ inaccentué par rapport au /i/ accentué est la suivante:

213%	125%	133%	225%	175%
------	------	------	------	------

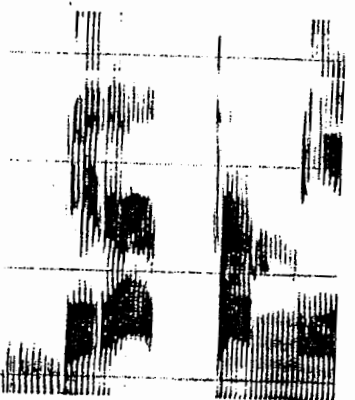
En ce qui concerne la constatation suivante citée dans des travaux différents, par ex. (2). "les consonnes se comportent tout-à-fait régulièrement dans l'entourage vocalique et les lois compliquées commencent à partir de groupes de consonnes", nous ne pouvons confirmer que la deuxième partie de la constatation mentionnant les "lois compliquées" (si l'on peut en effet parler des "lois"?)

On fait pour l'instant une observation assez importante dans le domaine des pauses pré-explosives. Le petit changement des pauses pré-explosives (une légère abréviation et même un léger allongement) se montre sur les images spectrographiques, par rapport aux autres éléments relativement stables. Les chiffres suivants présent (en %) la relation entre la première réalisation et les deux réalisations accélérées suivantes (résumé des réalisations

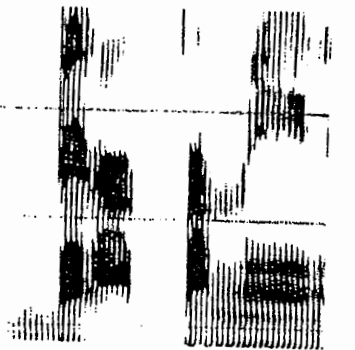
Comme illustration, nous présentons deux images spectrographiques



b i l a t a m ...

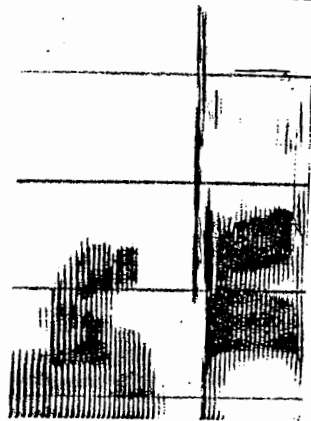


b i l a t a m ...

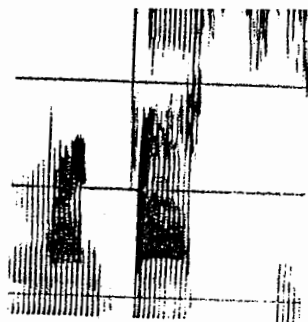


b i l a t a m ...

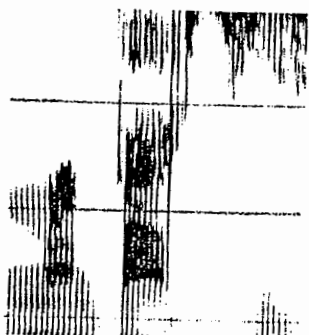
phiques reflétant l'augmentation de la vitesse du débit:



m o n t a :



m o n t a : š



m o n t a : š

de 5 locuteurs) - possibilité de comparer avec la durée des voyelles dans la même phrase:

a) Durée de la pause pré-explosive (en %).

/b/ 100,0	/t/ 100,0	/t/ 100,0
85,8	100,6	100,8
90,5	87,6	105,8

b) Durée de la voyelle /a/ (en%).

/a/ 100,0	/a/ 100,0	/a/ 100,0
85,8	78,0	87,0
80,0	71,0	73,0

c) Dans un autre phrase, les pauses pré-explosives sont les suivantes (/t/ par rapport à /a:/).

/t/	/a:/	/t/	/a:/
100,0	100,0	100,0	100,0
84,5	68,2	75,8	73,3
78,6	57,3	70,6	54,7

Après une analyse plus en profondeur des matériaux, nous arrivons à la constatation qu'il est préférable d'étudier toujours attentivement de petits ensembles en y examinant les phénomènes similaires ou contradictoires, attendus ou inattendus. Il est souvent très difficile de quantifier les mesures et les faits donnés, car de nombreuses compensations se manifestent ici, mais différemment chez plusieurs locuteurs, comme chez le même locuteur (dans les mêmes conditions).

Un travail statistique global effacerait les reliefs ainsi que les petits écarts contradictoires mais importants. Il est nécessaire de respecter une "microsegmentation" qui englobe aussi les phénomènes suivants: même dans la plus grande accélération on peut fixer sur les images spectrographiques (et ceci même en accord avec la perception) certaines unités stables, tandis que les autres unités vocaliques et consonantiques se groupent dans une "unité" nouvelle, ou elle perdent petit-à-petit leur stabilité et leurs contours.

Outre les observations concernant la stabilité relative des pauses pré-explosives, nous pouvons constater que cette recherche en cours apporte de nouvelles re-

marques surtout dans le domaine de la description quantitative des changements non linéaires pendant l'accélération du débit (de la parole). Les faits suivants peuvent encore servir d'illustration intéressante. Dans plusieurs réalisations (les exemples ci-dessus) se manifeste une augmentation de durée de certains éléments, même si la phrase au total (ainsi que ses segments) est continuellement abrégée. Par ex.: la réalisation la plus rapide de la 4^e phrase représente chez le locuteur n°1 une accélération de 46%, tandis que la pause pré-explosive devant le /t/ reste la même. Chez le locuteur n°2 l'accélération est de 36% et la pause s'allonge au contraire de 11%. Chez le locuteur n°3, l'abrègement est de 23% et l'allongement de la pause pré-explosive de 12,3%.

En conclusion, nous aimerions encore mentionner que les matériaux choisis ici apportent souvent toute une série de surprises, surtout en ce qui concerne les découvertes de nouvelles relations et connexions dont nous avons choisi au moins quelques exemples typiques. Ces problèmes peuvent être traités plus largement au cours de prochaines discussions.

REFERENCES

/1/ DOHALSKÁ-ZIHOVÁ, M. (1989), "To the testing of the dependence of transients on the speaking tempo", 28. akustická konference, Bratislava, 186-189.
 /2/ HUGGINS, A.W.P., (1972), "On the Perception of Temporal Phenomena in Speech", JASA, 51, 1271-1290.
 /3/ KLATT, D.H. (1976), "Linguistic uses of Segmental Duration in English", JASA, 59, 1208-1221.
 /4/ MALÁČ, V. (dir.) (1978), "Analýza časového členění věty", Tesla-VUST, Praha.
 /5/ UMEDA, N. (1977), "Consonant duration in American English", JASA, 61, 846-858.