

Effet du type d'activité cognitive et de nuisance auditive sur les caractéristiques du signal de parole

E. ABSIL et B. HARMEGNIES

Département de la Communication Parlée, Université de Mons-Hainaut, 18 Place du Parc, 7000 Mons, Belgium

Abstract : This paper reports acoustic analyses of 1. duration, 2. formant measures and 3. pitch performed on speech samples obtained in conditions of simultaneous presence of two independent variables : the type of cognitive task and the kind of auditive feedback. The results show modifications of the studied parameters and suggest an explanatory hypothesis about relation between tension, induced by stressing conditions, and larynx mobility.

1. INTRODUCTION

De nombreux travaux ont montré l'influence de divers types de nuisance auditive sur l'intelligibilité de la parole et sur les modifications acoustiques de celle-ci (1). Au niveau acoustique, les paramètres les plus couramment étudiés sont la fréquence fondamentale, l'intensité, la durée et les valeurs formantiques.

Le signal de parole est occasionnellement soumis à bien d'autres facteurs. Ainsi, quelques auteurs s'intéressent depuis peu à l'effet du stress sur la qualité vocale. On peut regrouper les études réalisées en laboratoire faisant appel à cette variable en trois types de méthode d'induction du stress : utilisation de stimuli nocifs (inhalation de solution amoniaquée), "modèle de supercherie" où les sujets sont sollicités pour mentir à l'expérimentateur, ou encore tâches de réalisation cognitive. En utilisant cette dernière stratégie, Griffin et William (4) ont déjà relevé une augmentation significative du fondamental et de l'amplitude, et une diminution de la durée des mots en fonction de l'accroissement de la complexité de la tâche.

La maîtrise expérimentale du type et du degré de nuisance est cependant éminemment variable de recherche à recherche. Par ailleurs, l'étude de certains facteurs perturbants - pour intéressante qu'elle soit au plan fondamental - ne semble pas ouvrir de perspectives réalistes dans le domaine de la recherche/développement. De ces points de vue, une investigation centrée sur les effets de la tâche cognitive paraît désirable, d'une part, eu égard aux possibilités de maîtrise aisée qu'offre cette dernière et d'autre part, en raison de la fréquente coexistence dans le chef du locuteur in situ entre activité cognitive spécifique et activité communicative (contrôle de processus commandés à la voix, surveillance du trafic pendant une communication téléphonique en voiture, etc.).

L'objet de la présente recherche est donc l'étude des répercussions sur le signal vocal du type de tâche cognitive, en interaction avec le type de nuisance auditive.

2. PLAN EXPERIMENTAL

2.1. Echantillon

Dix sujets masculins francophones, étudiant à l'Université de Mons-Hainaut ou y travaillant, ont été appelés à participer à l'expérience. Aucun de ceux-ci n'a auparavant subi de dommages phonatoires et/ou auditifs. Les sujets sont âgés de 20 à 47 ans.

2.2. Les variables indépendantes

La recherche présente tente de cerner les effets conjoints sur le signal de parole des variables indépendantes *tâche cognitive* et *nuisance auditive*.

La variable "*tâche cognitive*" présente deux modalités : "lecture simple" d'une part et "lecture/décodage", d'autre part. Dans les deux cas, un corpus invariant constitué de 40 mots français est produit. Chacun de ceux-ci comporte une des dix voyelles du français (soit /ɔ/, /oe/, /i/, /a/, /ɛ/, /y/, /e/, /ø/, /u/, /o/) en position préaccentuelle combinée à un environnement plosif ou fricatif non voisé. Pour chacun de ces deux modes articulatoires, les deux consonnes présentant les lieux d'articulation les plus extrêmes ont été retenues, soit /p/ (plosive la plus antérieure), /k/ (plosive la plus postérieure), /f/ (fricative la plus antérieure) et /ʃ/ (fricative la plus postérieure).

Pour la tâche de lecture simple, les mots sont insérés dans un texte continu narratif une historiette et d'une durée approximative de production d'une minute.

La production du corpus dans la seconde modalité s'obtient à la faveur d'une tâche d'association de mots à des nombres. Elle est présentée sous la forme de quatre feuilles d'exercices exploitant chacune 10 des mots du corpus. Chaque mot doit être produit à la faveur d'une association à un nombre-code. Lors de l'exécution de la tâche, le sujet dispose à sa gauche d'un écran sur lequel apparaissent les instructions ainsi qu'un sablier au moment de débiter la tâche. Celui-ci est programmé pour se remplir au fur et à mesure de l'activité mais de manière inconstante. De plus, des messages apparaissant à l'écran pressent le sujet d'activer son rythme d'exécution.

Les modalités de la deuxième variable indépendante sont 1° retour auditif en canal direct, 2° retour filtré passe-bas avec fréquence de coupure 800 Hz, 3° retour masqué par un bruit blanc à 92 dB filtré passe-haut avec fréquence de coupure 800 Hz. Les sujets testés perçoivent leur retour par le biais d'un casque tout au long de l'expérience.

Les différents niveaux de ces variables se croisent pour nous donner les six conditions d'étude des variables dépendantes que représente le tableau 1.

		NUISANCE AUDITIVE		
		Canal direct	Filtrage passe-bas	Masquage passe-haut
TACHE COGN.	Lecture simple	cond.1	cond.3	cond.5
	Lecture décodage	cond.2	cond.4	cond.6

Tabl.1 : structure du plan croisé à 2 variables indépendantes

2.3. Variables dépendantes

La durée, les formants et la fréquence fondamentale ont successivement été étudiés. Les deux premiers, portant sur l'analyse des voyelles, ont été relevés à l'aide du Sonographe digital Kay DSP 5500 dans la gamme DC-4 KHz. Le nombre de valeurs relevées est de 24000, soit 10 (sujets) X 6 (conditions) X 40 (voyelles) X 4 (paramètres). Le fondamental, quant à lui, a été mesuré, d'une part sur les voyelles, et d'autre part sur les mots, à l'aide du Kay Visi-Pitch 6095. Le nombre de valeurs relevées dans ces cas est de 160, soit 2 (sujets) X 2 (conditions) X 40 (voyelles ou mots).

Les informations subséquentes ne contiennent que les résultats globaux des traitements statistiques appliqués aux données recueillies étant donné les limites rédactionnelles à respecter.

3. ANALYSE DES RESULTATS

3.1. Durée

Nous avons procédé, pour évaluer les variations de durée (variable dépendante), à une analyse de variance croisant les critères "*tâche cognitive*" et "*nuisance auditive*" (variables indépendantes).

Cette analyse révèle des effets significatifs des variables indépendantes sur la durée des voyelles étudiées : $F(\text{tâche cognitive})=215.302$, $L=1$, $P(H_0)<.001$; $F(\text{retour aud.})=13.052$, $L=2$, $P(H_0)<.001$. L'effet est, par ailleurs, plus marqué en ce qui

concerne le type de tâche. Les valeurs moyennes des durées observées dans chaque condition indiquent une augmentation de cette durée avec la complexification de la tâche et suivant l'amplification de la nuisance auditive.

3.2. Formants

Pour chaque entourage consonantique, une analyse de variance des trois premiers formants selon les variables "retour" et "tâche" a été réalisée. Les traitements statistiques ont montré une significativité de F1 dans tous les cas, une non significativité généralisée de F2 et des résultats mitigés en ce qui concerne F3. La significativité est plus marquée au niveau de la variable "retour" pour le premier formant. Par contre, elle est plus importante au niveau de la variable "tâche" en ce qui concerne le troisième formant.

		F1		F3	
TACHE COGNITIVE	Lecture simple	m=379.35	σ =102.4	m=2456.81	σ =261.79
	Lecture/décod.	m=402.75	σ =100.36	m=2396.79	σ =311.68
NUISANCE AUDITIVE	Canal direct	m=379.47	σ =101.67	m=2442.66	σ =279.82
	Filtre P.B.	m=382.24	σ =94.34	m=2440.02	σ =289.2
	Bruit blanc	m=411.44	σ =106.74	m=2397.71	σ =296.85

Tabl.2 : moyennes et écarts types de F1 et F3 (Hz) selon les variables ind.

En observant les moyennes des premiers et troisièmes formants, nous constatons que la moyenne de F1 augmente avec la complexification de la tâche et l'amplification de la perturbation. La moyenne des F3, quant à elle, diminue avec la complexification de la tâche et tend à diminuer avec le niveau de perturbation du feedback.

Fry (5) et Laver (6) relient les variations de volume des cavités vocales et les modifications de la structure formantique.

A titre exploratoire, une hypothèse explicative inspirée de leurs raisonnements peut être avancée en ce qui concerne l'augmentation de F1 en condition 6 (tâche complexe et perturbation auditive maximale). Le stress accru en condition 6 pourrait être corrélé à une tension des cordes vocales, elle-même provoquant une élévation du larynx et donc une réduction de la cavité pharyngienne. Cette élévation laryngienne semble corrélée à l'augmentation de la fréquence de vibration des cordes vocales. Notre intérêt s'est donc ensuite porté sur les modifications éventuelles du pitch étant donné les variations observées de F1.

3.3. Pitch

Nous avons choisi de comparer les fréquences fondamentales de nos voyelles à l'intérieur des 40 mots produits dans chacune des conditions. Nous avons opté pour la confrontation des conditions 2 (tâche complexe sans nuisance auditive) et 6 (tâche complexe avec nuisance auditive) de l'expérience puisque c'est au niveau de cette condition 6 qu'est apparue l'augmentation maximale du premier formant. Cette analyse a été effectuée sur les deux sujets pour qui l'évolution de F1 était la plus différente (sujets 1 et 8).

Les tests t de Student pour échantillons appariés nous ont indiqué chez nos deux sujets des différences très significatives entre nos deux conditions au niveau de la moyenne de la fo, de sa valeur maximale et de sa valeur minimale.

Par contre, une non significativité des tests s'est révélée quant aux comparaisons des écarts types et des écarts entre les valeurs extrêmes.

Les significativités observées pour les trois premières statistiques sont, par ailleurs, toujours plus importantes chez le sujet 8 que chez le sujet 1.

Ayant montré qu'il y avait bien parallélisme entre les évolutions de F1 et du fondamental chez chacun de nos deux sujets, nous avons pu réaffirmer l'hypothèse explicative selon laquelle la condition de retour auditif masqué favoriserait, chez le sujet, une tension accrue, corrélée à une position élevée du larynx.

3.4. Dispersion

Bien que les indices de dispersion présentent une non significativité de leurs différences, des variations apparaissent entre les conditions 2 et 6 (C2 et C6). Lorsque, chez le sujet 1, F1 tend à augmenter entre C2 et C6, fo suit cette évolution. Nous notons, par ailleurs, une augmentation des indices de dispersion (écarts type et

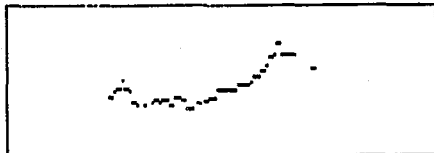
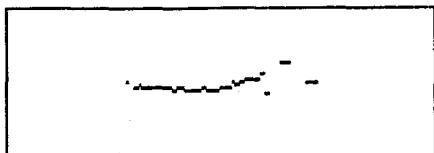
écarts entre les valeurs extrêmes).

Lorsque, chez le sujet 8, F1 tend à augmenter (plus encore), fo suit également cette évolution. Nous notons cette fois une légère diminution des indices de dispersion.

Pour tenter d'établir une relation entre les variations du fondamental et celles de la dispersion des données, nous avons décidé de comparer les patterns des 40 mots au point de vue du fondamental entre nos deux conditions.

L'analyse, dans ce cas, fut relativement subjective puisqu'il s'agissait d'observer et de comparer les patterns d'un même mot produit dans deux conditions différentes.

Nous avons caractérisé nos patterns de "plus dispersé en condition 2", "de plus dispersé en condition 6", ou enfin ayant une "dispersion plus ou moins semblable".



Mot "puni" : pattern de fo en cond.2 (S1)

Mot "puni" : pattern de fo en cond.2 (S1)

Il semble que la dispersion qui a été observée au niveau des voyelles se retrouve au niveau de ces patterns.

Dispersion	C2 > C6	C6 > C2	C2 = C6
Sujet 1	2	28	10
Sujet 8	18	5	17

Tabl.3 : dispersion des patterns des mots (au niveau du fondamental) entre les conditions 2 et 6 pour les sujets 1 et 8

En effet, pour le sujet 1, nous remarquons une dispersion majoritairement plus importante en condition 6, là où nous avions noté une augmentation des écarts.

Pour le sujet 8, les résultats sont moins tranchés puisque nous trouvons autant de patterns plus dispersés en condition 2 que de patterns n'évoluant pas. La dispersion moins importante des patterns en condition 6 rejoint la petite diminution observée au niveau des écarts.

La dispersion de la fréquence fondamentale serait donc fonction de sa valeur moyenne. Il apparaît, qu'en position élevée, le larynx ait moins de possibilité à se mouvoir; c'est ce que nous avons noté chez le sujet 8.

De même, la moindre augmentation de fo chez le sujet 1 permet à ce dernier une importante liberté de déplacement du larynx, établie ici par sa propension à varier son pattern.

4. CONCLUSION

L'ensemble des résultats nous permet de conclure à des effets de nos variables indépendantes sur la plupart des variables dépendantes étudiées. L'altération du retour auditif et le type de tâche cognitive produisent, de manière générale, des modifications sur les caractéristiques du signal de parole, et notamment sur la durée des voyelles, sur leurs premiers et troisièmes formants ainsi que sur leur fondamental.

5. BIBLIOGRAPHIE

(1) Van Summers W., Pisoni D.B., Bernacki R.H., Pedlow R.I., Stokes M.A., " Effects of noise on speech production : Acoustics and perceptual analyses ", *Journal of Acoustical Society of America*, 84(3), 1988, 917-928.
 (2) Lindblom B., " Spectrographic study of vowel reduction ", *Journal of Acoustical Society of America*, 35(11), 1963, 1773-1781.
 (3) Delattre P. in Durand P., " Variabilité acoustique et invariance en français, consonnes occlusives et voyelles ", (Editions du CNRS, collection "Sons et parole", 1985, Paris).
 (4) Griffin G.R., Williams C.E., " The effects of different levels of task complexity on three vocal measures ", *Aviation, space and environmental medicine*, 58, 1978, 1167-1170.
 (5) Fry D.B., " The physics of speech ", (Cambridge Textbooks in Linguistics, Cambridge University Press, 1979).
 (6) Laver J., " The phonetic description of voice quality ", (Cambridge University Press, 1980, Cambridge).