

Effets de la compression temporelle d'indices acoustiques sur la perception de la parole

C. Jacquier & F. Meunier

Laboratoire Dynamique Du Langage. CNRS UMR 5596, Institut des Sciences de l'Homme
14, avenue Berthelot. 69363 LYON CEDEX 07 – FRANCE.

Contacts: jacquier@isc.cnrs.fr – Fanny.Meunier@univ-lyon2.fr

Contexte

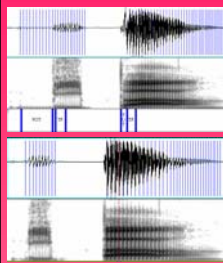
Notre étude explore les mécanismes cognitifs impliqués dans la reconstruction de la parole dégradée. Nous nous intéressons aux caractéristiques temporelles d'indices acoustiques spécifiques utilisés par les auditeurs pour distinguer les sons de parole. Pour distinguer deux phonèmes, un intervalle de temps minimum entre deux stimuli est nécessaire afin de pouvoir traiter séparément les deux signaux sans les confondre. De nombreuses études ont établi que le déficit de traitement phonologique chez des dyslexiques serait le reflet d'un déficit de traitement et d'intégration du signal de parole lors de changements rapides de l'information. Selon cette hypothèse, cette incapacité à analyser des transitions brèves du signal empêche de discriminer les indices acoustiques nécessaires à la distinction des phonèmes (Paula Tallal, 1980). Dans notre étude nous avons manipulé ces contraintes temporelles entre différents indices acoustiques afin de tester leur pertinence dans l'identification de pseudomots et d'identifier les mécanismes de reconstruction du signal de parole à différents niveaux (phonétique et acoustique).

Dans notre étude, nous avons compressé sur l'axe temporel à la fois le Voice Onset Time (VOT) et la transition formantique. Le VOT est défini comme l'intervalle de temps entre le début du burst de la consonne et le début du voisement de la voyelle. Le VOT est un bon indice temporel pour discriminer sur le plan acoustique des consonnes voisées et non voisées. La transition formantique, entre le F2 de la voyelle et le locus de la consonne, permet de distinguer deux plosives par rapport à leur lieu d'articulation.

Matériel et Méthode

Participants

- 32 volontaires : 10 M et 22 F (âge moyen = 24.44 ans) / langue maternelle : Français sans trouble du langage, ni trouble auditif



Oscillogramme et Spectrogramme du pseudomote [bipa] avec la segmentation des VOT et des transitions formantiques (TF) dans la condition initiale 100% (en haut) et dans la condition 0% pour les deux syllabes (en bas).

Stimuli

- Occlusives Voisées /b/ et /d/
- Occlusives Non Voisées /p/ et /t/
- Voyelles /a/ et /i/
- ⇒ 64 pseudomots bisyllabiques de forme C1V1C2V2
- ⇒ 16 fillers de forme V1C1V2
- Enregistrement dans chambre sourde par locuteur masculin

Méthode

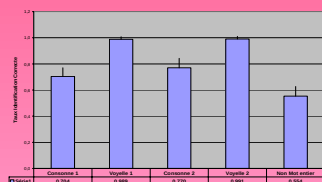
- Segmentation des VOT et des transitions formantiques
- 4 conditions d'accélération :
 - 100% = 100% de la durée initiale VOT et transition
 - 50% = reste 50% de la durée initiale VOT et transition
 - 25% = reste 25% de la durée initiale VOT et transition
 - 0% = déletion complète VOT et transition

Tâche : Identification auditive des pseudomots avec transcription au clavier.

Partie I :

Effets de la compression sur les mécanismes cognitifs de reconstruction des phonèmes

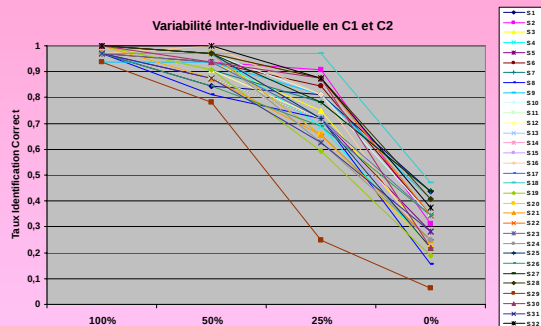
- V1 et V2 sont identifiées quasiment à 100%
 - validation du paradigme expérimental
- C2 est mieux identifié que C1
 - allongement de consonne intervocalique
- Identification Totale de plus de 55% des non mots



- La perte d'identification de C2 intervient à partir de la condition 25% alors que la perte de C1 apparaît dès 50%.

- La voyelle /i/ est mieux identifiée que la voyelle /a/.

- Il existe une grande variabilité inter-individuelle en C1 et C2 : l'écart type le plus important est pour la condition 25%.



Perspectives : batterie de tests de lecture et de tests audiométriques pour tenter de corréler les performances des sujets à des contraintes linguistiques ou auditives.

Explication de cette grande variabilité observée à partir d'un seuil spécifique de dégradation acoustique chez des sujets sains.

Conclusion

L'ensemble de ces résultats met en évidence:

- Une hétérogénéité des performances inter-individuelles.
- Une hétérogénéité des performances des consonnes occlusives selon leur position dans le pseudomote, selon leur contexte vocalique et selon le taux de compression des indices acoustiques.

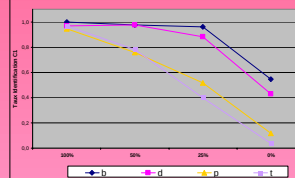
Maintenant, il reste à tester la compression des indices acoustiques séparément afin de déterminer l'implication de chacun dans les effets observés.

Partie II :

Effets de la compression sur l'identification des traits acoustiques spécifiques des phonèmes

En position 1 :

- Les consonnes voisées (/b/, /d/) sont mieux identifiées que les non voisées (/p/, /t/).
- La perte d'identification des voisins n'apparaît qu'à partir de la condition 25% alors que la perte pour les non voisées apparaît immédiatement à 50%.
- De plus, les labiales (/b/, /p/) semblent toujours mieux identifiées que les dentales (/d/, /t/).

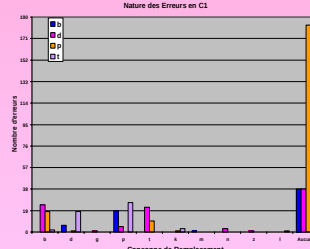


En position 2 :

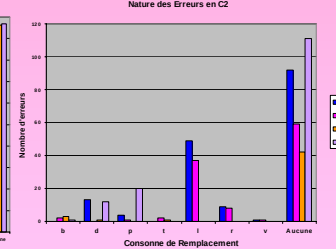
- Globalement, les non voisées sont mieux identifiées que les voisées.
- Les deux labiales /p/ et /b/ montrent une très forte différence d'identification, au profit du /p/ qui est quasiment à 100% d'identification.



En position 1 : les non voisées /p/ et /t/ sont majoritairement perdues dès 50% de compression alors que le peu d'erreurs sur les voisées sont des erreurs de lieu d'articulation ou de voisement.



En position 2 : les voisées /b/ et /d/ dès 50% de compression sont soit perdues soit confondues avec des consonnes liquides /l/ et /r/, la majorité des confusion sont observées avec un contexte vocalique /i/ (V2).



Perspectives : descriptions acoustiques des items (durée du VOT, trajectoire et pente des transitions) afin de corréler la qualité d'identification des sujets à des caractéristiques acoustiques et phonétiques.

Démonstration de l'importance des indices acoustiques testés dans l'identification de signaux de parole naturelle.