



HAL
open science

Articulation silencieuse vs. vocalisée pour une ”communication parlée silencieuse” : implications des différences articulatoires

Cédric Gendrot, Lise Crevier-Buchman, Bruce Denby, Thomas Hueber,
Antonia Colazo-Simon, Claire Pillot-Loiseau, Pierre Roussel, Sophie
Quattrocchi

► **To cite this version:**

Cédric Gendrot, Lise Crevier-Buchman, Bruce Denby, Thomas Hueber, Antonia Colazo-Simon, et al..
Articulation silencieuse vs. vocalisée pour une ”communication parlée silencieuse” : implications des
différences articulatoires. JPC 2011 - 4èmes Journées de Phonétique Clinique, May 2011, Strasbourg,
France. pp.29-30. halshs-00610689

HAL Id: halshs-00610689

<https://shs.hal.science/halshs-00610689>

Submitted on 23 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Articulation silencieuse vs. vocalisée pour une « communication parlée silencieuse » : implications des différences articulatoires.

Cédric Gendrot², Lise Crevier-Buchman², Bruce Denby^{3,1}, Thomas Hueber⁴, Antonia Colazo-Simon², Claire Pillot-Loiseau², Pierre Roussel¹, Sophie Quattrocchi².

¹ SIGMA Laboratory, ESPCI ParisTech, CNRS-UMR 7084, Paris, France

² Laboratoire de Phonétique et Phonologie, CNRS-UMR 7018, Paris, France

³ Université Pierre et Marie Curie, Paris, France

⁴ GIPSA-Lab, Département Parole & Cognition, CNRS-UMR 5216, Grenoble, France

cgendrot@univ-paris3.fr

L'idée d'une communication parlée silencieuse a récemment reçu une attention considérable de la part de la communauté de parole [1]. Ces outils qui effectuent une reconnaissance de la parole depuis des capteurs appliqués contre le conduit vocal, sont conçus afin de permettre à leur utilisateur de communiquer en silence (un « téléphone silencieux » par exemple), ou afin de prévoir une alternative à la voix trachéo-œsophagienne pour les personnes ayant perdu l'usage de la parole suite à une laryngectomie. Ces systèmes doivent être le moins invasifs possibles, et dans une étude récente [2-4], nous avons proposé un système portable d'analyse en temps réel des mouvements de la langue et des lèvres, à partir respectivement d'un appareil à ultra-sons et d'une caméra.

Des études effectuées en anglais sur un système semblable [3-4] bien que fixe, ont permis d'aboutir à des scores de 70% de taux de reconnaissance. En dehors des problèmes matériels de conception d'un outil maximale portable, les données utilisées pour ces études consistaient en des phrases vocalisées, ce qui ne correspond pas aux attentes d'un outil qui vise à être appliqué à des patients laryngectomisés. En effet, des différences significatives ont pu être observées entre la parole silencieuse et la parole vocalisée, notamment à cause du manque de retour auditif. Dans la présente étude, nous décrivons des tests de reconnaissance de mots produits à la fois en parole silencieuse et en parole vocalisée dans les phases de tests et d'entraînement. Dans un premier temps [6], un système d'acquisition fixe a été comparé à une ébauche de système portable avec des scores de reconnaissance très similaires, permettant de valider l'efficacité d'un système portable, et ainsi d'envisager l'utilisation dans un avenir proche de tels systèmes pour utilisateurs handicapés.

Les résultats du système de reconnaissance à base de modèles de Markov [6] ont mis en avant une part beaucoup plus importante de la langue - comparée aux lèvres - dans le score final. Malgré tout, les modèles de Markov ne permettent pas de décrire les mouvements des lèvres et de la langue effectués par les informateurs. Nous montrons ici des différences quantifiées d'aperture, d'étirement et d'arrondissement des lèvres, ainsi que des mouvements d'inflexion, de hauteur et d'antériorité de la langue pour la parole silencieuse comparée à la parole vocalisée. Les résultats montrent pour la parole vocalisée des stratégies similaires au renforcement des traits distinctifs des phonèmes pour la parole hyperarticulée, à savoir des phonèmes arrondis plus arrondis, des voyelles ouvertes plus ouvertes, etc. dans le cas de la parole vocalisée.

En [4], nous mettons en avant la nécessité d'entraîner le système de reconnaissance sur des productions silencieuses. Cette étude permet d'en comprendre les raisons et pourquoi pas, de mieux les appréhender. Le travail final présentera également une discussion du type de corpus à utiliser et de la méthodologie des interfaces de communication parlée silencieuse

This work was supported by the French National Research Agency (ANR) under contract number ANR-09-ETEC-005-01 REVOIX.

Denby, B., Schultz, T., Honda, K., Hueber, T., Gilbert, J.M. and Brumberg, J.S., "Silent speech interfaces", *Speech Communication*, 52, pp. 270-287, 2009.

Hueber, T., Benaroya, E.L., Chollet, G., Denby, B., Dreyfus, G. and Stone, M., "Development of a Silent Speech Interface Driven by Ultrasound and Optical Images of the Tongue and Lips", *Speech Communication*, 52, pp. 288-300, 2009.

Hueber, T., Chollet, G., Denby, B., Dreyfus, G., and Stone, M., "Visuo-Phonetic Decoding using Multi-Stream and Context-Dependent Models for an Ultrasound-based Silent Speech Interface," in *Interspeech*, Brighton, UK, pp. 640-643, 2009.

Florescu, V-M., Crevier-Buchman, L., Denby, B., Hueber, T., Colazo-Simon, A., Pillot-Loiseau, C., Roussel, P. Gendrot, C., Quattrocchi, S. (2010), "Silent vs Vocalized Articulation for a Portable Ultrasound-Based Silent Speech Interface", *Proceedings of Interspeech* (Makuri, Japan).