Umeå University, Department of Philosophy and Linguistics
PHONUM 9 (2003), *89-92*
Available online at http://www.ling.umu.se/fonetik2003/

Testing a new method for training fricatives using visual maps in the Ortho-Logo-Paedia project (OLP)

Anne-Marie Öster[1], David House[1] Athanassios Hatzis[2], and Phil Green[2]

[1]*Department of Speech, Music and Hearing, KTH*

[2]*Department of Computer Science, University of Sheffield*

This paper presents new techniques used to produce visual maps for training Swedish fricatives for hearing-impaired clients, and presents results from testing one such map with the speech of three hearing-impaired subjects. The results demonstrate the utility of the mapping techniques for visually portraying correct and deviant pronunciations and indicate the potential for use in actual clinical practice. The study is carried out within OLP, an EU-funded research project which makes use of real-time acoustic-to-articulation kinematic mapping and speech recognition techniques for speech training.

1. Introduction

The OLP (Ortho-Logo-Paedia) project aims at applying a method to supplement (but not replace) speech therapy for specific disorders at the articulation level based on an integrated computer-based system together with automatic speech recognition and distance learning. The key features of the project are: (a) therapy based on real-time audio-visual feedback of client's speech; (b) speech production evaluation and interfacing to assistive technology provided through automatic speech recognition; and (c) web services to provide remote collaboration and data collection for analysis and evaluation in diverse conditions.

This paper concentrates on the first feature (a) and presents results from building and testing visual maps for training hearing-impaired clients using the OPTACIA component. OPTACIA was developed from the Optical Logo-Therapy OLT (Hatzis, 1999; Hatzis and Green, 2001). OPTACIA is based on three basic, well-founded treatment principles: visuomotor tracking, visual contrast, and visual reinforcement. Visuomotor tracking (Ziegler, Vogel, Teiwes, and Ahrndt, 1997) is a special case of biofeedback where some dynamic physical measure of performance is portrayed visually in real-time. Visual contrast is a contrast between the normal, the misarticulated, and the produced sound pattern to build client awareness of differences in various articulatory configurations (Öster, 1996). Visual reinforcement in general can be considered as a reinforcer when it increases the rate of response or some quality of it. The aim of the therapy with OPTACIA is to increase, strengthen, and eventually maintain the frequency of correct and accurate responses.

2. Swedish fricative articulation

The most common articulatory deviation by Swedish normal-hearing and hard-of-hearing children is an incorrect pronunciation of the Swedish sibilant fricatives (represented here by Sampa symbols) /s/ [s], /C/ [ç] and /S/ [ɧ]. Sibilant fricatives are those with a higher frequency and greater acoustic energy than non-sibilant fricatives like /v/, /f/, /h/ and /j/. The difficulty to perceive and produce these speech sounds by profoundly hearing-impaired persons is partly due to the fact that their hearing losses usually involve higher frequency ranges where the acoustic properties of these speech sounds are situated. From an articulatory point of view these sounds cause difficulties, as the precise formation of the constriction as well as the accurate place of the constriction is critical for correct realisation. The airflow must not be too weak, a fact that implies a correct physiological control. A common deviation by profoundly hearing-impaired speakers is that /s/, which requires a passage of air through a groovelike opening, is substituted by a lisping sound produced with a flat placement of the tongue tip between the upper and lower teeth. In some instances /s/ is also replaced by /C/ or /S/ because of a somewhat back position.

3. Building the Swedish fricative map

To produce the map, it was necessary to have training data, in the form of recorded speech time-annotated wave-files with target labels of those speech sounds, which are to appear on the map. The design of the Swedish sibilant fricatives map with acoustic and phonetic information of the target sounds is shown in Figure 1 (left panel).

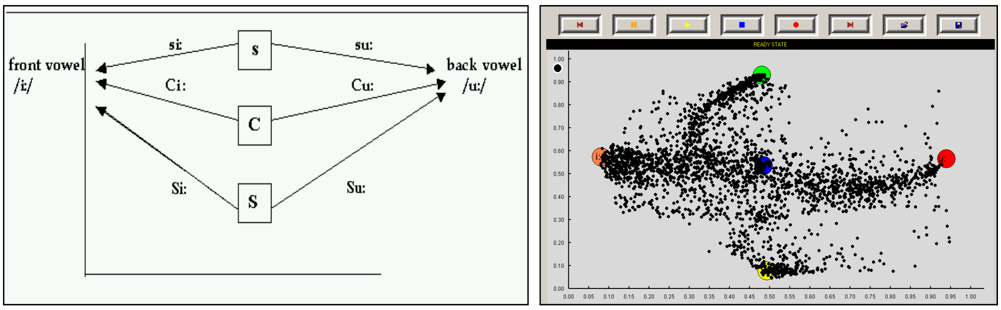


Figure 1. Design of the map for Swedish training of the sibilant fricatives (left) and layout of the actual Swedish sibilant fricative map with trial data (right), both in Sampa symbols

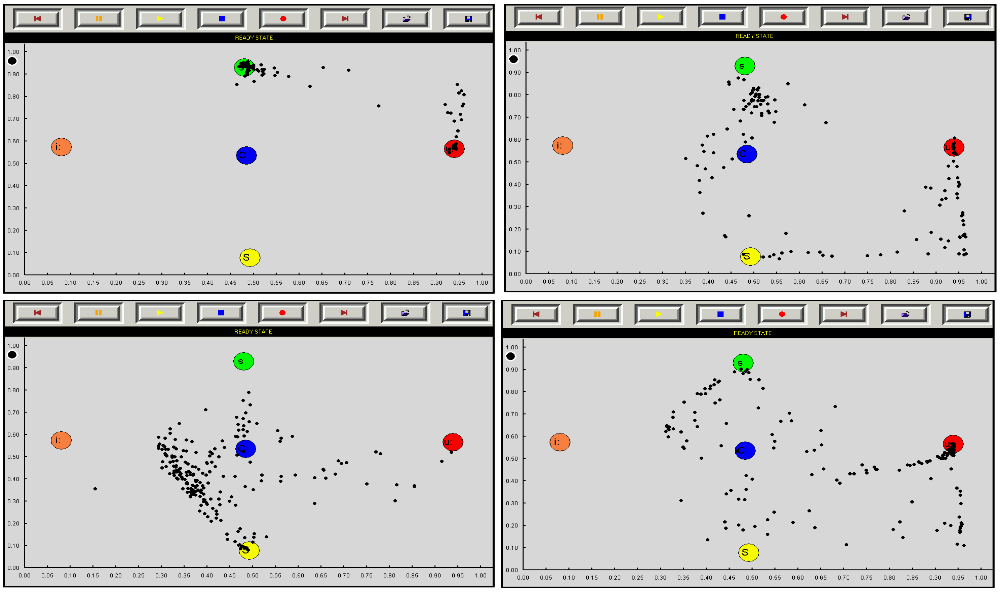
The acoustic difference between the three Swedish sibilants is represented by the vertical axis, which shows frequency range. By the positions of the fixed points along the axis the relationship between the acoustics and the articulation of each Swedish sibilant fricative is demonstrated. Articulatory targets for training each sibilant together with either a front spread /i/ or a back rounded /u/ in syllables are also inserted in the map providing visual paths during training.

Recordings were made of a 12 year-old normal-hearing girl when she pronounced repetitions of CV utterances where C was a sibilant fricative /s/, /C/ and /S/ and V was either /i/ or /u/. Each CV combination was recorded five times. Each recording contained an average of six repetitions. The data files were recorded and manually labelled in the Wavesurfer application, http://www.speech.kth.se/wavesurfer/. The data files were then subjected to a cepstral analysis, and the resulting labelled 12-component vector formed the training input to a multi-layer perceptron (MLP), which has two outputs, corresponding to a 2D position of the targets according to the design of the map.

4. Testing the map

As an initial live-input test, the map was presented to a 17 year-old normal-hearing female control subject. She was instructed to first say the phonemes in isolation one by one to get acquainted with the system. She was then instructed to pronounce all six CV combinations to chart the visual articulatory trajectories. The resulting trajectories are shown in Figure 1 (right panel). All five phoneme targets are well represented on the map as are the trajectories from all fricatives to /i/ and /C/ to /u/. However, the trajectories from /s/ and /S/ to /u/ are not as clear.

The recorded speech of three profoundly hearing-impaired children served as a test-material for the map. The children (one girl and two boys) were 16 years of age and belonged to a school for the deaf in Stockholm. The children were recorded when they repeated CVC combinations. All the produced combinations of /sis/ [si:s] and /sus/ [su:s] were selected and played back to the map.



Figures 2. The result of two hearing-impaired children's repetitions of /sis/ and /sus/ played back to the map. The left panels are /sis/ and the right panels are /sus/. The upper panels are from the girl and the lower ones from the boy.

The results of two of the children (one girl and one boy) can be seen in Figure 2. The upper left panel shows the girl's production of /sis/. In this production, she is quite successful at producing the fricative, but the vowel maps onto the /u/ instead of the intended /i/. This could be the result of an incorrect /i/ production which includes lip-rounding and strong nasalization. In the upper right panel, the same girl has intended a pronunciation of /sus/ but here the map indicates the instability of her fricative. Her /u/ vowel is correctly pronounced.

In the lower left panel, the result of the boy's /sis/ pronunciation is shown. He has more difficulty with the fricative /s/ than the girl, but he comes closer to an acceptable pronunciation of the /i/ vowel. Finally, in the lower right panel, the result of the boy's /sus/ is shown. In this attempt, his first /s/ is quite successful, as is his vowel. It is the /s/ in final position which he has difficulty with. The results for the other boy were not as clear. As he had low intelligibility and a very indistinct articulation, the map showed no targets or trajectories.

5. Conclusions

The results of the initial testing of the Swedish fricative map for showing visual feedback with OPTACIA demonstrate a promising potential for use in training articulation of hearing-impaired clients. An auditory analysis of the two children's deviant pronunciations of the CVC combinations matches very well with the acoustic articulatory positions on the visual map. This means that the system is powerful enough to be used with clients who produce deviant hearing-impaired speech. However, to avoid the problem of interspeaker variability, the client's best productions should be stored and used to retrain the map. The new map is then used for further therapy to establish automaticity and maintenance. This possibility of pronunciation training with visual comparison is an effective strategy, used in several visual aids. Furthermore, the possibility in OPTACIA to easily change the layout or to visualize articulatory movements rather than sounds, e.g. tongue retraction, nasality, lip rounding, etc, seems to be unique and promising.

6. Acknowledgements

The OLP project http://www.xanthi.ilsp.gr/olp/ is coordinated by the Institute for Language and Speech Processing, Athens, Greece, with participation from France (Arches), Greece (Altec S.A. and Logos Centre for Speech-Voice Pathology), Spain (Universidad Politécnica de Madrid), Sweden (KTH), and the UK (Sheffield University and Barnsley District General Hospital). The project is funded by the European Commission, Fifth Framework Programme, Quality of Life and Management of Living Resources, contract QLG5-CT-2001-01971.

7. References

Hatzis, A., (1999) *Optical Logo-Therapy (OLT), Computer-based audio-visual feedback using interactive visual displays for speech training*, Department of Computer Science, University of Sheffield, UK.

Hatzis, A. and Green, P. D., (2001) A two dimensional kinematic mapping between speech and acoustics and vocal tract configurations, *Workshop on Innovation in Speech Processing, (WISP'01)*, Stratford-upon-Avon, UK, 2001.

Ziegler, W., Vogel, M., Teiwes, J., and Ahrndt, T. (1997) Microcomputer-Based Experimentation, Assessment and Treatment, In Ball, M.J., and Code, C. (Eds), *Instrumental Clinical Phonetics*, Whurr Publishers.

Öster, A-M., (1996) Clinical applications of computer-based speech training for children with hearing impairment, *Proceedings of ICSLP 96*, 157-160, Philadelphia, USA.