USING THE CANADIAN SPEECH RESEARCH ENVIRONMENT (CSRE) TO TEACH SPEECH PERCEPTION TO UNDERGRADUATES

Margaret F. Cheesman and Donald G. Jamieson
Hearing Health Care Research Unit
Department of Communicative Disorders
University of Western Ontario
London, ON N6G 1H1

This paper describes four computer-based learning (CBL) sessions that were designed for and used in the speech acoustics and perception sections of a course in Speech Science offered by the Department of Communicative Disorders at the University of Western Ontario. Students enrolled in the course have undergraduate degrees in diverse disciplines, including the physical, social, and biological sciences, arts and humanities. Consequently, students vary greatly in their previous experiences with audio and computer hardware, computer software packages, and in their knowledge of speech perception and physical and speech acoustics.

All of the laboratory experiences described below were developed with the Canadian Speech Research Environment (CSRE; Jamieson, Ramji, Kheirallah and Nearey, 1992). CSRE is a comprehensive, microcomputer-based system designed to support speech research using IBM/AT-compatible microcomputers. CSRE provides a powerful, low-cost facility to support speech research using mass-produced and widely-available hardware. The software has been used in several graduate and undergraduate courses since 1989. Cheesman and Alsop (1990) reported on initial experiences with CBL modules used with this system. The most recently released version, CSRE4.0[1], is currently being used for CBL applications. The software includes speech sampling, editing and replay functions; a selection of frequency analysis procedures and colour two- and three-dimensional displays; formant and other parameter extraction and tracking; parametric speech (KLATT80) and non-speech synthesizers, and a comprehensive experiment generator and controller that support behavioural testing using several common testing protocols.

For each of the laboratory sessions, students are given a printed handout containing basic instructions and guidelines for the laboratory session. Students prepare a written report, following a set of directed questions included on the handout.

1. Analysis of misarticulated speech sounds

In this session, students digitally record and spectrographically analyze their own productions of normal and misarticulated speech sounds that they have encountered in their clinical phonetics course. The lab assignment is structured as a step-by-step sequence that results in a tutorial introduction to the use of the CSRE sampling, editing and analysis functions. Figure 1 illustrates a spectrographic comparison, using auto-correlation, of a normal and misarticulated "play".

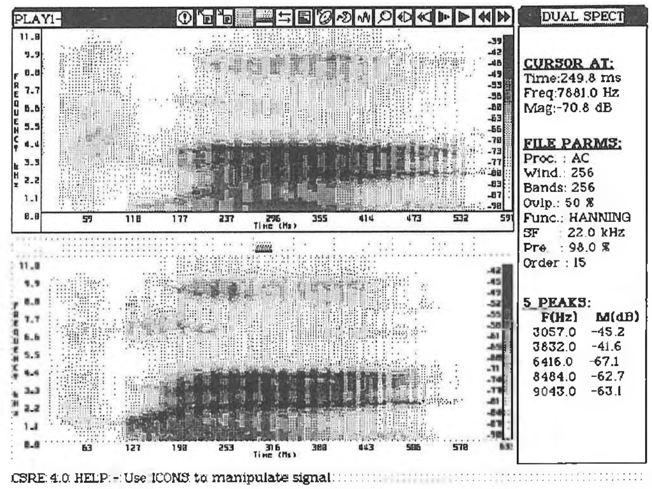

Figure 1. Dual spectrogram of normal and misarticulated /plei/

2. Vowel perception laboratory

This three-part laboratory experience focuses on the relationship between formant frequencies and vowel perception and allows students to test their own perception. Acoustical analysis of test items allows students to confirm the observations they have made in the perceptual testing.

I. Effect of varying formants 1 and 2. A two-dimensional grid of F1 and F2 space is presented. Each point on the grid corresponds to a synthetic vowel which the student can listener to repeatedly. Students transcribe each of the sounds and then examine the pattern of transcription with the formant space. In the example in Figure 2, the sound corresponds to an /ɛ/-like vowel with a first formant centred at 550 Hz and a second formant at 1650 Hz.

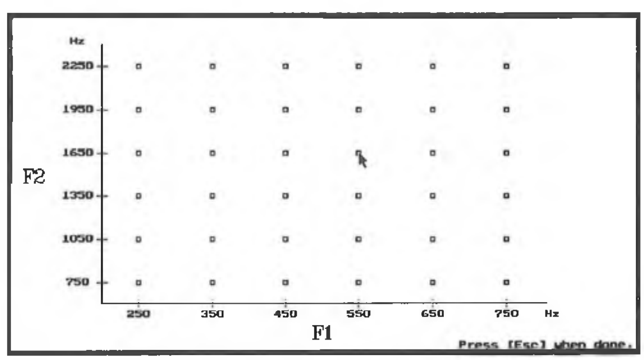

Figure 2. Example of monitor display during vowel perception module.

II. Non-English vowel sounds. A series of foreign-language (German) vowel sounds, spoken in various phonetic contexts, are presented. Students select one of four vowel categories to assign to each vowel and are given correct answer feedback. The task is repeated three times, so that students can examine their initial performance and any learning effects. The exercise underscores the dominant role of the native language categories in the perception of even relatively steady-state foreign contrasts.

III. Acoustical analysis of vowel sounds. The German vowels used in the task described above are analyzed with respect to formant frequencies. Performance on the F1/F2 vowel space and German vowel identification tasks are re-analyzed using this information.

3. Stop/semi-vowel perception laboratory

I. Acoustic and perceptual analysis of consonants. Students analyze, display and listen to the medial consonant portion of four digitized nonsense words taken from the modified distinctive features difference test (DFD[m] - Cheesman, Lawrence and Appleyard, 1992). Using the partial playback feature of CSRE, students can listen to and attenuate selected segments of the words in order to study the perceptual consequences of removing selected acoustic information.

II. Creation of stop consonant percept. Using the editing features of CSRE, students insert a silent interval after the fricative portion of the nonsense word /ʌslI/ to observe the perceptual effect. This silence is typically perceived as the oral closure associated with stop consonants and the dominant percept is /ʌsplI/. Figure 3 shows the waveform editing process used to isolate the end of the frication noise.

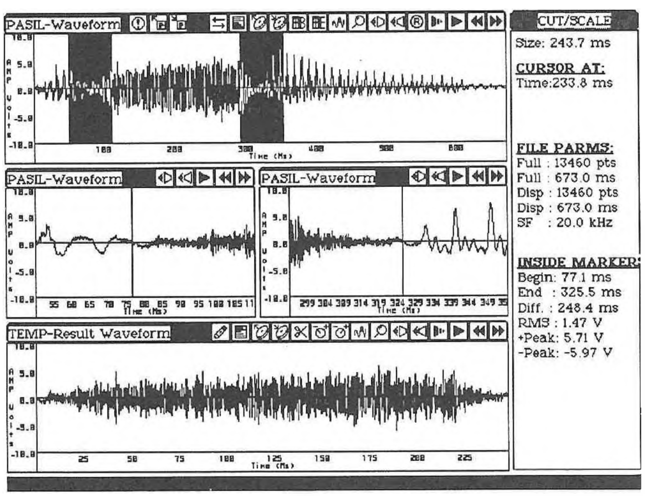

Figure 3. Waveform display of the nonsense word /ʌslI/ during editing.

4. Fricative perception and context effects

I. Context-conditioned voiceless fricatives. Synthetic stimuli consisting of nine bandpass noises followed by /i/ and /u/ vowels are identified by students as /s/ or /ʃ/ and the effect of the vowel on the identification of the fricative continuum is examined. Students also compare their own data with that of other class members to observe individual differences in category boundaries and context effects.

II. Comparison of synthetic speech acoustics to natural tokens. Spectrographic analyses of natural and synthetic /s/ and /ʃ/ sounds are made and the results are compared. Figure 4 displays such a comparison. Students are encouraged to examine alternative

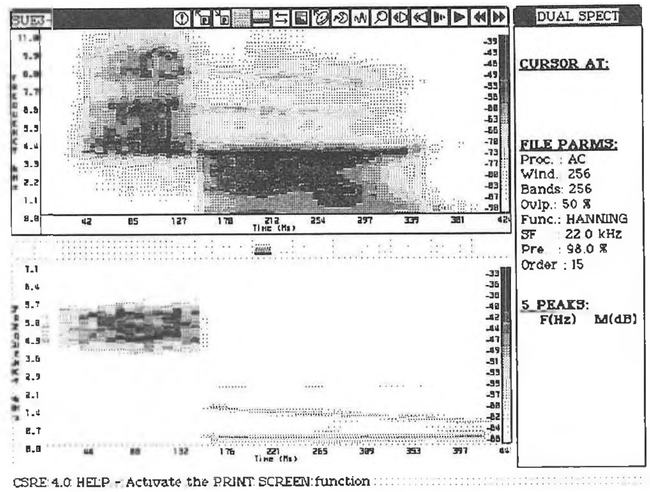

Figure 4. Spectrograms of naturally-produced and synthetic tokens of /su/.

spectral representations of the same signal. For example, they may compare the representation obtained with traditional linear predictive analysis to the cone-kernel spectrogram.

Acknowledgements

Thanks to W. Alsop and K. Ramji for assistance in the development of the CBL software. P. Kirchberger assisted with preparation of the report. This research is supported by NSERC.

[1]CSRE4.0 software is available from the Hearing Health Care Research Unit at the University of Western Ontario.

References

Cheesman, M.F. and Alsop, L.W. (1990). Computer-enhanced teaching of speech acoustics. **Proceedings of Acoustics Week in Canada - 1990**, 41-45.

Cheesman, M.F., Lawrence, S., and Appleyard, A. (1992). Prediction of performance on a nonsense syllable test using the Articulation Index. In Ohala, J. J., Nearey, T. M., Derwing, B. L., Hodge, M. M., and Wiebe, G. E. (Eds.), **ICSLP 92 Proceedings: 1992 International Conference on Spoken Language Processing**. Edmonton: University of Alberta. pp. 1123-1126.

Jamieson, D.G., Ramji, K., Kheirallah, I., and Nearey, T.M. (1992) CSRE: The Canadian speech research environment. In J.J. Ohala, T.M. Nearey, B.L. Derwing, M.M. Hodge, and G.E. Wiebe (Eds.) **Proceedings of the Second International Conference on Spoken Language Processing**, Edmonton: University of Alberta, pp. 1127-1130.