TEACHING PHONETICS USING THE PHONETIC DATA BASE ON MICROCOMPUTER

JOHN H. ESLING

Department of Linguistics
University of Victoria
Victoria, B.C. V8W 2Y2 Canada

ABSTRACT

Development of a microcomputer-based speech processing, analysis, manipulation and input/output system allows the incorporation of revised techniques in the teaching of basic phonetics. The Micro Speech Lab and related speech editing software permit acquisition, storage, random-access retrieval, variable-order selection, marking, concatenation, and auditory and visual comparison of phonetic data. Phonemic inventories and speech samples of several diverse languages illustrating a variety of phonetic constraints have been collected for research and instructional applications in a Phonetic Data Base. Tasks are described which give phonetics students the opportunity to collect speech sound data, hear and evaluate linguistic and indexical contrasts, and extract short samples for illustration, comparison and practice.

The MICRO SPEECH LAB SYSTEM

The procedures for phonetics instruction described here are the direct result of the development of a Micro Speech Lab microcomputer-based system for capture, playback and analysis of speech and other acoustic signals. Micro Speech Lab (MSL) is a complete hardware/software package for use with IBM PC, XT, AT microcomputers, designed and developed in the Centre for Speech Technology Research/Phonetics Laboratory at the University of Victoria. MSL contains a software diskette or hard disk and documented on paper by number digitized using the MSL capturing routine on the IBM-PC microcomputer. Files are stored by language as diacritic or hard disk and documented on pages by number for reference to phonetic, phonemic and orthographic representations and English gloss of each sample. Examples of phonetic sounds that are normally difficult to obtain, and phonemic inventories of a range of languages not usually encountered or available during the course of phonetics classes have been included. Languages collected thus far include: Egyptian Arabic, Inuktitut, Korean, Miriam, Nittinat, Nungangsams, Rutroo, Scots Gaelic, Skagit (Coast Salish), Spokane, Turkish, Umpug, Xhosa, and Yoruba. At least 50 words and several short text files for each language have been stored in the current library.

This system has been made available to students in phonetics classes, including those in Applied Linguistics (Teaching English as a Second Language) teacher preparation programs. Individual words and short texts are accessed from diacritic or hard disk directories and manipulated by groups of students according to tasks set by the instructor on particular auditory categories. MSL packages are currently available with MSL EDIT, and must be performed outside of class time.

REVISING METHODS OF INSTRUCTION

The Phonetic Data Base (PDB) is intended to provide a practical, accessible and realistic mechanism for experimenting in a practice-oriented approach to the study of phonetics and phonology. The PDB gives manipulative and creative power to learners, allowing them to test hypotheses, reorganize elements in new ways, and critically evaluate the results of their work. The goal is to enhance phonetics instruction by incorporating the PDB and MSL delivery system with a number of new task-oriented activities that will fit into the structure of the phonetics course. Revised techniques emphasize the role of prosody, including voice setting, in the initial stages of phonetic exposure rather than focusing exclusively on the auditory analysis. Attention is given to the interpretation of indexical as well as phonological properties of speech, for listeners first exposed to a new linguistic and cultural environment. Auditory recognition and assignment of written symbols to represent categories of sounds are the central skills to be developed here. Laboratory sessions in the second language (L2) listening/speaking tasks that emphasize aural discrimination rather than production [6] [7]. Many current, popular L2 learning programs incorporate listening training of pronunciation [8] [9] and may leave language teachers with no clear model of how to present L2 speech sounds other than their recollection of how they were taught phonetics. This decrease in overt attention given to the pronunciation component of L2 teaching has caused some alarm [10], and it is hoped that this discrepancy can be partially reduced by introducing a modified approach into the course where language teachers originally learn phonetics.

The emphasis in L2 teaching is shifting away from the static model approach based on the ideal phonemic inventory of the target language taught in a dedicated pronunciation class, towards communicative, problem-solving task-oriented activities designed to provide a larger context of L2 manipulation for language students [11]. Where pronunciation is taught explicitly in L2 programs, the focus has shifted to word-level meaning contrasts rather than phoneme drills, and to the early introduction of prosodic features [12] [13] [14] [15]. Specific conditions found to benefit L2 acquisition include a range of language learning strategies and a variety of listening/speaking tasks, assimilated at the student's own pace, (4) presence of significant target language models, especially of peers [16]. The benefit for phonetics teaching is the range of 'dynamic output' that can be presented with the random-access capability of the system for comparing and evaluating the range of auditory features. This is illustrated in figure 2 where the Inuktitut text has been analyzed to show amplitude (middle screen) and pitch (bottom screen) over time. Read-outs represent values at the position of the left cursor. In figure 2, left and right cursors have been placed to isolate 25 frames of speech. This adjustable window, or the entire waveform, can be monitored using D/A by pressing the function keys indicated in the menu scroll. This capability is also present in MSL EDIT.

Pitch, amplitude and spectral characteristics can also be calculated and displayed by MSL, adding recognition of visual correlates to the task of becoming familiar with a range of auditory features. This is illustrated in figure 2 where the Inuktitut text has been analyzed to show amplitude (middle screen) and pitch (bottom screen) over time. Read-outs represent values at the position of the left cursor. In figure 2, left and right cursors have been placed to isolate 25 frames of speech. This adjustable window, or the entire waveform, can be monitored using D/A by pressing the function keys indicated in the menu scroll. This capability is also present in MSL EDIT.

Students locate new sounds, manipulate elements of stored items, arrange them in categories, and create new sets of files that reflect the inventory of speech sounds from the phonetic chart to meet the instructional objective of...
the course. Evaluation of the task considers (a) number and range of representation of sounds collected, (b) adequacy of each item extracted from surrounding speech to illustrate the sound intended, and (c) organization of items into phonetic categories for presentation. The goal is for students to become active agents in their own learning process while, at the same time, learning the use of instrumental techniques.

Items are collected with task-based instructions: "Find all the words from the following languages that have sound X in them," and then "Find the sounds that the following words have in common," and "Group the following words together according to sounds that they share in common." Once collected, the sounds are studied in detail and gradually assigned phonetic symbols. Sounds isolated in this process are then grouped together in new files representing sets of allophonic variants of the "same" phoneme in a language. Isolated sounds can also be combined in new files represented by the same phonetic symbol, but which have been taken from different languages. Figure 3 illustrates how short samples can be collected, marked and displayed. The cursor in each screen is aligned at 0.021 sec to highlight initial consonant differences.

**Figure 3.**

MSL EDIT display of similar CV sequences.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xhosa: [f]</td>
<td>Korean: [t]</td>
<td>Nithi: [s]</td>
<td>Skagit: [t]</td>
</tr>
</tbody>
</table>

If pharyngeal sounds need to be demonstrated, for example, the pharyngealized series of stops, affricates or nasals from Salishan and Wakashan languages, or the pharyngealized series from Arabic, are loaded for auditory contrast and transcription and visual observation of acoustic correlates. Extensive exposure is achieved by having students collect a variety of reflexes of each articulation specified on the phonetic chart, especially for sounds or symbols they find difficult and want to practice.

In another activity format, as a testing or "challenge" procedure, five items are displayed for visual identification. The instructor or a student specifies a sound by phonetic symbol or articulatory label, for a group of students to locate. Cursors can be positioned on the screens to isolate the sound and examine its transitions. If the indicated sound is not present, the item(s) closest to it in articulatory features must be identified.

**CONCLUSION**

With the development of a Phonetic Data Base, the presentation of speech sound material for phonetic study is facilitated, allowing expedient access to greater amounts of data, and manipulation and organization of speech items in an active learning format. The system also permits the training of language teachers in the use of technological aids for the delivery of speech sound information, in a manner consistent with the precepts of communicative, holistic language learning theory. Research on second language acquisition processes and teaching approaches is integrated with Micro Speech Lab hardware and software for delivery and analysis of speech signals to provide an expedient system for presenting phonetic material for pedagogical purposes. Additional applications of this system include the transmission and sharing of speech data for collaboration in phonetic research.

**REFERENCES**


Acknowledgements. I am indebted to my colleagues, Geoff O'Grady, Thom Hess, Barry Carson, and numerous colleagues and graduate students who were the source of the majority of linguistic material in the Phonetic Data Base. The indispensable cooperation and collaboration of Craig Dickson and Roy Snell of the Centre for Speech Technology Research at the University of Victoria in conceiving and developing microcomputer-based speech processing systems is gratefully acknowledged, as are general university research grants from the Social Sciences and Humanities Research Council of Canada.