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 contrasts 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 and internally mounted data acquisition hardware including anti-aliasing filters, A/D and D/A circuitry, and a user's manual with user instructions and descriptions of theory of use and applications [1]. The software includes user control of signal input, several waveform displays, audio output, analysis (amplitude, pitch, spectrum) and file management. Phonetic instruction using MSL applies this rapid random-access speech input/output capability to the recording, storage, recall, comparison, visual waveform observation and manipulation, spectral analysis, and variable auditory presentation of speech sound material collected in a Phonetic Data Base.

In addition, a program written to supplement MSL's speech-capturing, storage and processing capabilities, MSL EDIT, allows students to access and display graphic waveforms of sampled data files in order to listen to words or several-second samples of text in any language selected, vary listening sequences, edit existing files, and combine elements of old files into new files. "Designed as a supplementary package to accompany the Micro Speech Lab, the purpose of the program is to provide a highly flexible method for auditory examination and manipulation of digitally stored signals" [2]. Up to five sampled data files can be displayed and monitored individually, in reverse, or in continuous repetition of sequences composed of parts of any file.

THE PHONETIC DATA BASE

To provide a core of linguistically organized speech data for instructional and research purposes, a Phonetic Data Base of speech samples has been assembled using MSL. Words and text drawn from numerous linguistic, sociolinguistic and dialect survey sources, represent a wide range of speech sounds of languages of the world. Samples are digitally encoded using the MSL capturing routine on the IBM-PC microcomputer. Files are stored by language on diskette or hard disk and documented on paper 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 most phonetics classes have been included. Languages collected thus far include: Egyptian Arabic, Inuktitut, Korean, Miriam, Nitinaht, Nyangumarta, Rutooro, Scots Gaelic, Skagit (Coast Salish), Spokane, Turkish, Umpila, 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 diskette or hard disk directories and manipulated by groups of students according to tasks set by the instructor to focus on particular auditory categories. Most tasks are carried out with MSL EDIT, and most are 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 experiencing, comparing and evaluating the range of the multi-dimensional acoustic space reflected in the phonetic chart. The PDB gives manipulative and creative power to learners [3], and illustrates to prospective teachers how students can be enabled to collect and store language items in a format that allows easy reorganization [4] [5].

The goal is to enhance phonetics instruction by incorporating the PDB and MSL delivery system with a number of recent developments in second language acquisition theory 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 attention immediately on segmental analysis. Attention is given to the interpretation of indexical as well as of phonological properties of speech, for listeners first encountering a new language.

Auditory recognition and assignment of written symbols to represent categories of sounds are the central skills to be developed here, as they are in second language (L2) listening/speaking tasks that emphasize aural discrimination rather than production [6] [7]. Many current, popular L2 teaching approaches omit explicit teaching 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 themselves 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.

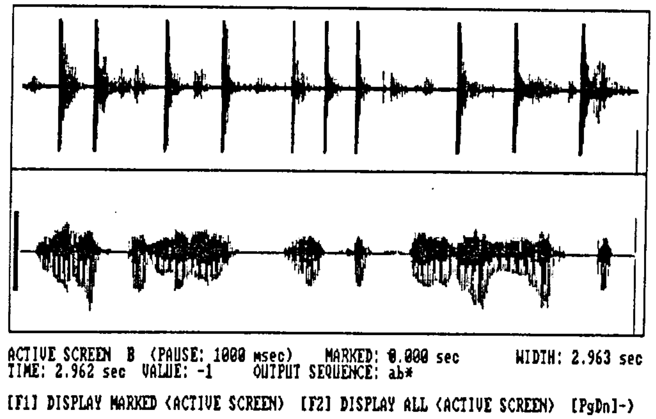
The emphasis in L2 teaching is shifting away from the static model approach based on the ideal phonemic inventory of the target language taught in a dedicated pronunciation class, towards communicative, problem-solving task-based activities designed to provide larger amounts of L2 for manipulation by 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 and teaching include: (1) diversity of language material presented in meaningful situations, (2) experience and practice in perception before production is required, (3) clear identification and association of concrete referents, 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 variety" that can be presented with the random-access capability of MSL. Instead of relying on imitation and memorization in production of segmental units, MSL allows students to satisfy perceptual criteria by obtaining understandable "input" of sounds in their natural context before having to produce "output" [17]. Students can choose their own pace and sequence using MSL, rather than having recorded examples presented in uncontrolled order. The criterion of identification with speaker models is met as students use MSL to choose specific individuals' voices to listen to and arrange spoken material into phonetic classification schemes [18]. In this way, students gradually build an inventory of sounds and symbols, to complete the phonetic chart, based on "input" which they have collected and classified themselves according to the phonetic principles they are learning.

TASK-ORIENTED INSTRUCTION IN BASIC PHONETICS

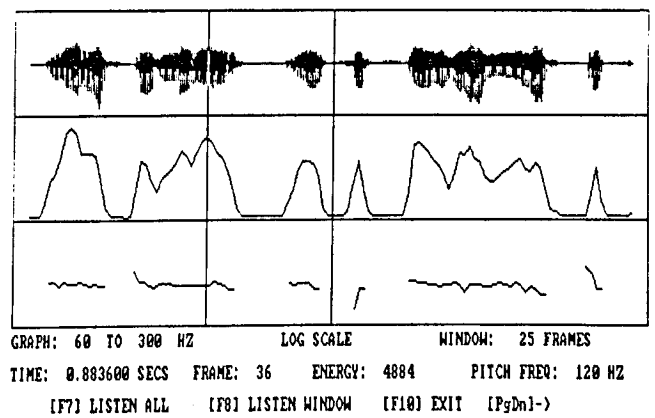
To use the PDB, groups of 2-4 students sit at a "work-station" table around one computer with external speaker and space for writing and reference materials. All words and texts are listed by filename, including language and number, to be typed in when calling up an item. Lists include phonetic and phonemic transcriptions of each item, orthographic representations in the native language if available, and English glosses. The instructor first assigns types of sounds for students to listen for, with lists of files containing those sounds. Initial listening focuses attention at the long-term level, on features of phonation type, voice setting, and dynamics [19]. Some features which can be identified from PDB text samples in this long-term listening exercise include tongue fronting, breathiness, close jaw, backing, spreading, roundedness, nasality, and prominent manners or secondary articulations such as retroflexion, clicks, approximants, frication, affrication, or glottalization. Figure 1 illustrates an MSL EDIT display contrasting recurring clicks of Xhosa (screen A, top) with the acoustic waveform and articulatory characteristics of an Inuktitut text (screen B). The set-up indicates that screens A and B will be heard repeatedly at 1000msec intervals.

Figure 1.
MSL EDIT display of Xhosa (top) and Inuktitut text.



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.

Figure 2.
MSL amplitude and pitch display of Inuktitut text.

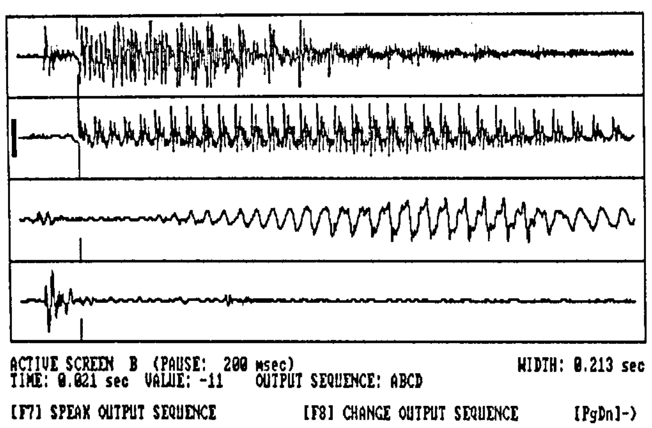


Students locate new sounds, manipulate elements of stored items, arrange them in categories, and create new sets of files that represent the inventory of speech sounds from the phonetic chart to meet the instructional objective of the course. Evaluation of the task considers: (a) number and range of 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 0.021sec to highlight initial consonant differences.

Figure 3.
MSL EDIT display of similar CV sequences.

A	B	C	D
Xhosa [t'ɑ]	Korean [t+ɑ]	Nitinaht [t'ə]	Skagit [t'ə]



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

[1] C. Dickson, *User's manual for Micro Speech Lab*, Software Research Corporation, 1985.
[2] C. Dickson, *User's manual for the MSL comparison and editing program: MSL EDIT*, Centre for Speech Technology Research Society, 1987.
[3] F. Smith, *Reading without nonsense*, Teachers College Press, 1978.
[4] E. Stevick, *Memory, meaning and method*, Newbury House, 1976.
[5] S. Savignon, *Communicative competence: theory and classroom practice*, Addison-Wesley, 1983.
[6] J. Morley, *Improving aural comprehension*, University of Michigan Press, 1972.
[7] P. Dunkel, Developing listening fluency in L2: Theoretical principles and pedagogical considerations, *Modern Language Journal, 70*, 99-106, 1986.
[8] S. Krashen, T. Terrell, *The natural approach*, Pergamon, 1983.
[9] E. Purcell, R. Suter, Predictors of pronunciation accuracy: A reexamination, *Language Learning, 30*, 271-287, 1980.
[10] R. Wong, Does pronunciation teaching have a place in the communicative classroom? Georgetown University Round Table on Languages and Linguistics, June, 1985.
[11] G. Brown, G. Yule, *Teaching the spoken language*, Cambridge, 1983.
[12] H. Woods, *Syllable stress and unstress*, Canadian Government Publishing Centre, 1978.
[13] J. Esling, R. Wong, Voice quality settings and the teaching of pronunciation, *TESOL Quarterly, 17*, 89-95, 1983.
[14] J. Gilbert, *Clear speech*, Cambridge, 1984.
[15] B. Harmegnies, A. Landercy, Language features in the long-term average spectrum, *Revue de Phonétique Appliquée, 73-74-75*, 69-79, 1985.
[16] H. Dulay, M. Burt, S. Krashen, *Language two*, Oxford, 1982.
[17] S. Krashen, *Second language acquisition and second language learning*, Pergamon, 1981.
[18] T. Terrell, Acquisition in the natural approach: The binding/access framework, *Modern Language Journal, 70*, 213-227, 1986.
[19] D. Abercrombie, *Elements of general phonetics*, Edinburgh University Press, 1967.

Acknowledgements. I am indebted to my co-researchers, Geoff O'Grady, Thom Hess, Barry Carlson, and numerous colleagues and graduate students who are 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.