

Teaching Tone and Intonation With Microcomputers

Dorothy M. Chun
University of Texas at Austin

ABSTRACT: Although research on the use and effectiveness of visual feedback for teaching tone and intonation began more than thirty years ago, the technology for signal analysis and pitch extraction using microcomputers has only recently become widely accessible and affordable. This paper 1) reviews the major pedagogical applications of acoustic phonetic research for teaching segmentals (individual sounds) and suprasegmentals (intonation, stress, rhythm); 2) summarizes the hardware and software currently available for speech analysis on the Macintosh and IBM-PCs, and 3) discusses courseware features that should be included in implementing this new technology to help foreign language students improve their production and perception of tone and intonation.

KEYWORDS: teaching tone and intonation (suprasegmentals), visual feedback, signal analysis on the Macintosh and IBM-PCs, sound digitizers, spectrograms, pitch-tracking software, courseware development.

Introduction

This paper deals with the use of microcomputers, in particular, the Macintosh, to provide visual feedback for students learning individual tones in languages such as Chinese or for those trying to learn more native-like intonation in languages such as German, French or English.¹ Although research on the use of visual feedback in teaching tone and intonation began over three decades ago (e.g., Vardanian 1964, Abberton & Fourcin 1975, James 1976 & 1979), the technology for signal analysis and pitch extraction using microcomputers has only recently become widely accessible and affordable. First, I briefly review previous research on the effectiveness of visual feedback in teaching pronunciation and the pedagogical applications of acoustic linguistic research and technology to date. The second part of the paper describes the hardware and software currently available for the Macintosh and the IBM personal computers, with particular focus on inexpensive sound digitizers and pitch-tracking programs for the Macintosh. To date, there are no widely used programs for the

teaching either of individual sounds or of tone and intonation, but the availability of the hardware and software described herein should facilitate development of pronunciation software. The final section of the paper thus discusses courseware features that should be included in the implementation of this new technology to help foreign language students improve both their production and their perception of tone and intonation.

Application of Speech Technology to Teaching Pronunciation

One of the fastest growing areas of research in computer technology is speech recognition, i.e., getting computers to recognize and "understand" human speech. This current focus has led to the commercial development and production of voice-activated machinery. Fortunately, it has also spawned some practical applications for the teaching of pronunciation. Hardware and software for teaching both segmentals (i.e., individual sounds) as well as suprasegmentals (i.e., intonation, stress and rhythm) are emerging. Molholt (1988), for example, reports on four years of research in the application of voice-activated machinery to language education. He notes that traditional methods of correcting students' pronunciation rely heavily on subjective evaluations, e.g., teachers saying "No, no that's not right, say it like this," or students having to recognize their own errors in a language lab exercise by comparing their pronunciation to that of a native speaker on a master tape. The result is that many errors go undetected and become fossilized. A computer display, however, of pronunciation comparing a native speaker's model with students' attempts to match it, can give students "objective information about the location, extent, type, and significance of the error, as well as the progress made in correcting the error" (p. 92). In the following three subsections, we look first at research on the effectiveness of visual feedback and then at specific examples of teaching both individual sounds (segmentals) as well as sentence melody or intonation (suprasegmentals). The examples will demonstrate the strengths of the various display types for learners.

A. Effectiveness of Visual Feedback

Lon and Martin (1972, 141) suggested that visualization of the intonation pattern allows for the "establishment of an automatic judgment system based on theories of pattern recognition." Experiments by James (1979) provided evidence that visualization can have a significant effect on improving the intonation of second language learners. Three groups of subjects were observed: the first group followed a traditional approach of listening to and repeating model sentences; the second group was given an instantaneous visual representation of the intonation contour for each model sentence, but no feedback of their own repetitions; the third group saw immediate visualization of both the model sentences as well as their own imitation. The second group did not perform

significantly better than the first; i.e., the method providing visualization of the model resulted in little or no improvement over the traditional method of imitating an auditory model. However, the third group, which received immediate reinforcement in the form of visual feedback, was far superior to the other two groups.

Research by de Bot (1980 & 1983) showed that audio-visual feedback is more effective in intonation learning than auditory feedback. In the 1983 study, the two factors in the experiment were feedback mode and practice time. The results showed a significant effect of audio-visual feedback over auditory feedback, whereas practice time did not seem to be a major factor. In other words, optimum imitation of a sentence was reached sooner with audio-visual feedback than with auditory feedback only. The type of feedback provided also influenced learning behavior: those in the audio-visual feedback groups repeated target sentences more often than subjects with auditory feedback only, whereas the latter listened more often to their own imitations than the former. This suggests that subjects with audio-visual feedback decided mainly on the basis of what they saw whether to repeat an old sentence or to start a new one. The auditory feedback groups could only make this decision on the basis of what they heard, and since they received no external information about their learning success, they may have been less motivated to practice, repeat, or try to correct an error.

B. Teaching Segmentals

An example of teaching individual sounds can be found in Molholt (1988, 92-3). In teaching aspiration of a stop like /p/ in initial position in English words, a spectrographic display of the correct sound as well as of the sound uttered by a student who is not accustomed to aspirating the /p/ provides concrete and objective feedback showing the differences between the two sounds. Figure 1 displays the spectral frequencies of the speech signals [pht], [ptl]

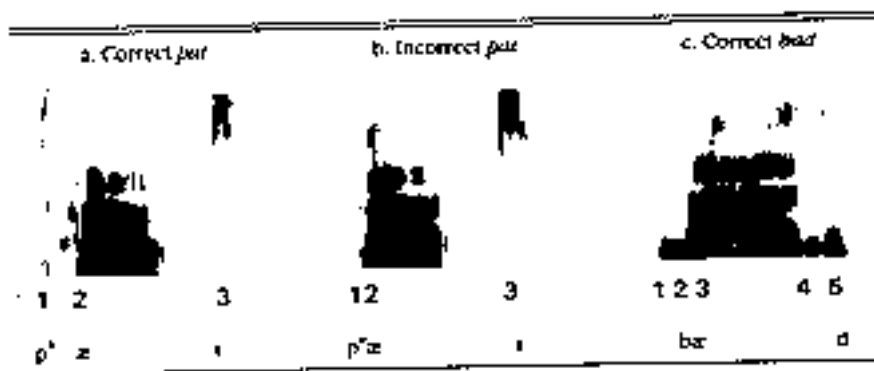


Figure 1: SSD Displays

and [bd] made by a Speech Spectrographic Display (SSD) 8800. (The vertical axis is frequency in Hz; the horizontal axis is duration; and the third dimension, amplitude, is indicated by the darkness of the display, i.e., the darker the display the stronger the signal.) Such displays can show a student instantly and clearly that, e.g., in Figure 1b, there was not enough of a break and therefore not enough aspiration between the initial explosion of the /p/ (point 1) and the beginning of the vowel sound (point 2). This lenis [p] can be contrasted with Figure 1a, the aspirated fortis [ph], which clearly has a greater distance between points 1 and 2. (Figure 1c shows the voiced [b] as well as voiced [d], the short distance between the initial explosion of the /b/ (at point 2) and the following vowel (point 3), and the longer vowel which is used before voiced consonants in English.) In actuality, the SSD provides a split-screen display so instructors' and students' sounds may be compared. This sort of "bio-feedback" for pronunciation helps accelerate the acquisition of correct pronunciation "because the visual display provides an objective measure that helps students focus their attention on the exact features that need to be changed" (Molholt 1988, 96).

The problem with aspiration is common among French and Spanish speakers learning English, with the reverse difficulty experienced by English speakers learning French and Spanish, to name two of the most widely taught languages. Spectrograms are not only useful for teaching consonants but perhaps, more importantly, for teaching vowels. Common problems for American learners are the nasal vowels of French and the front rounded high and mid vowels (the umlauts) of both French and German. Spectrographic feedback would easily allow students to compare their attempts with that of native speakers. One obvious drawback, however, is that students must first be told what causes the differences between their pronunciation and the model native speaker's and must be able to recognize the source of their error (the source of the "mismatch") before they can actively work on changing their output.

C. Teaching Suprasegmentals

As with the individual sounds or segmentals, spectrograms can provide useful information for teaching suprasegmentals, particularly with regard to pitch or frequency, duration, and amplitude. (Recall that in the spectrograms in Figure 1, the vertical axis is frequency, the horizontal axis is duration, and the third dimension, the darkness of the display, is amplitude.) However, this information is not as easily read or interpreted from spectrograms as information on voicing or individual vowels. Thus, other types of displays are preferable, e.g., those which extract fundamental frequency (Fo) information from a digitized speech signal and subsequently plot the contour on a display.

Figure 2 (also from Molholt 1988, 94) displays the same three words as Figure I but was made on a Visi-Pitch 6095 (described in more detail below). Like the spectrographic displays, displays of speech signals are provided instantly but are much simpler: they provide information only about the strength and pitch of the lowest frequency range (0-400 Hz). The lower line corresponds to voicing and pitch (or fundamental frequency), the upper to the relative strength or intensity of all the sounds. Thus, since the top curve starts farther to the left than the bottom line, there is an interval of silence (or voicelessness) from the beginning of /p/ to the vowel /ae/. The cursor has been placed at the onset of the vowel to

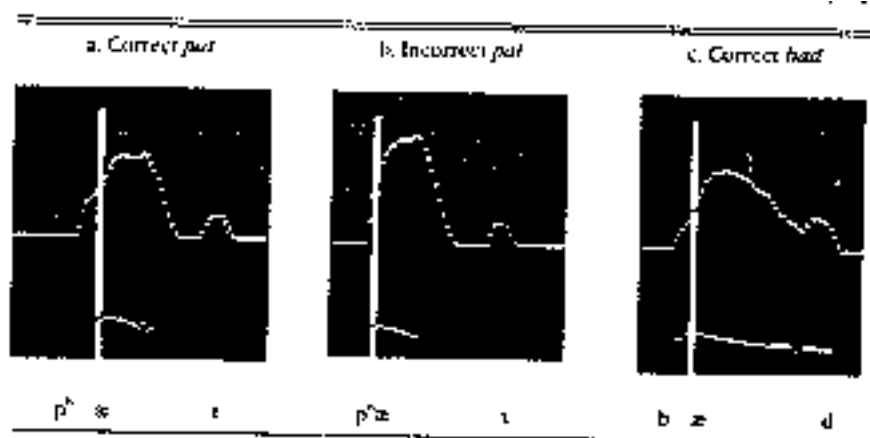
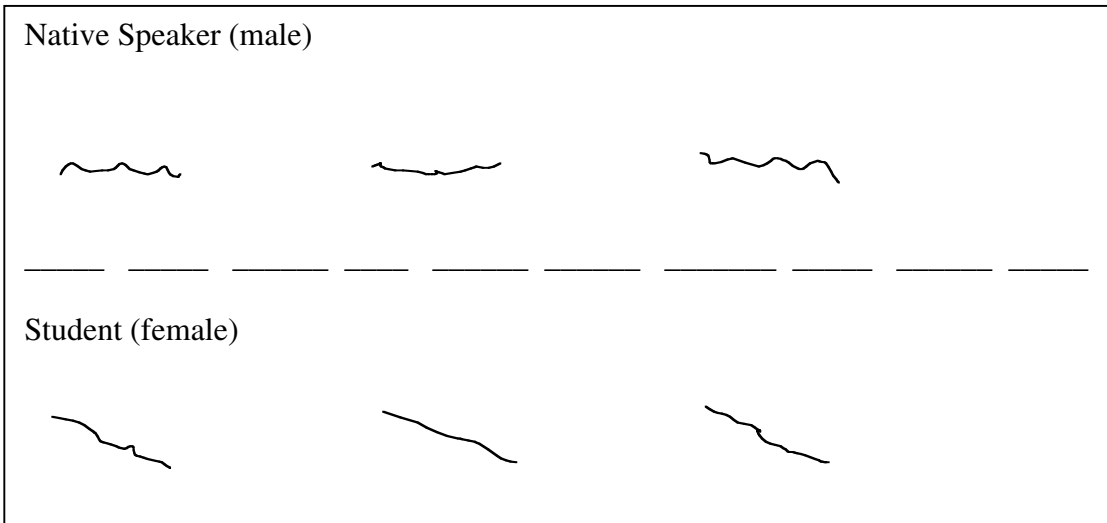


Figure 2: *Visi-Pitch Displays*

show this clearly. Figure 2b shows the incorrect version corresponding to Figure 1b. The onset of the vowel is too close to the beginning of the word, so listeners will think the word is bat rather than pat. One also notes that the pitch of the voice falls slightly for each of the three words and that bad is much longer in duration than the other two. (pp. 93-94)

The display of fundamental frequency (perceived as pitch or tones) and, to a somewhat lesser extent, amplitude (perceived as intensity or loudness) are particularly useful for teaching tone languages like Chinese. Figures 3 and 4 (from Fischer 1986a) show Visi-Pitch displays of the pitch of three repetitions of two monosyllabic words in Mandarin Chinese as spoken by a native speaker (upper curves) and a learner (lower curves). Tone I is a level tone (Figure 3), but is produced by the student with slightly falling pitch. Tone 4 is a falling tone (Figure 4); the student pronounces it as falling but also adds a little dip before the peak and uses more of a concave falling contour than the native speaker. (Note also the statistics available to the learner, e.g., column #1 shows the native speaker's statistics, column #2 the student's; values are given for average Fo

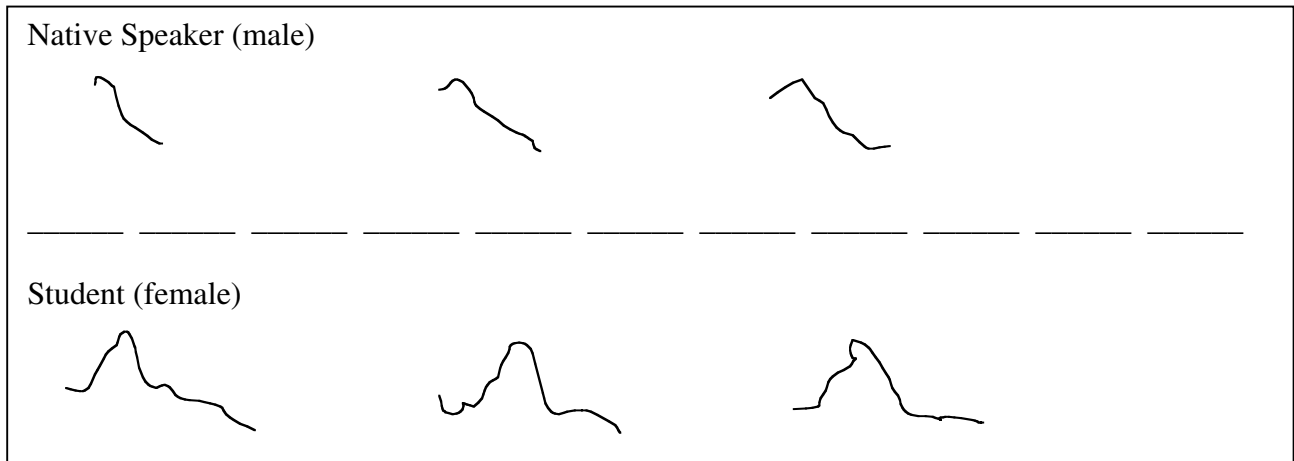
(pitch), Fo maxima and minima (range), and duration.) Figure 5 shows both intensity and pitch curves for three repetitions of a disyllabic word *wenti* ("questions") which consists of tone 4 followed by tone 2 (a rising tone) and was meant as a question ("Questions?"). Comparison of the pitch contours shows that the student fails to use a falling 4th tone for the first syllable of the word and uses a rise for the 2nd tone which is longer and rises higher than the native speaker's. In addition, examination of the intensity curves is also instructive: the student uses greater intensity for the second syllable than for the first whereas the native speaker does the reverse.



Values for all data between cursors

STATISTIC	COLUMN #1	COLUMN #2	CHANGE
AverageFo	133.1	222.4	89.3 Hz
Time between cursors	000.464	000.752	000.288 sec
Maximum Fo	141.5	257.4	115.9 Hz
Minimum Fo	125.9	191.4	065.5 Hz
Fo Range	015.6	066.0	050.4 Hz

Figure 3: Visi-Pitch Displays Chinese Tone 1



Values are for all data between cursors

STATISTIC	COLUMN #1	COLUMN #2	CHANGE
Average Fo	118.9	172.6	053.7 Hz
Time between cursors	000.432	000.640	000.208 sec
Maximum Fo	161.7	233.9	072.2 Hz
Minimum Fo	070.1	141.1	071.0 Hz
Fo Range	091.6	092.8	001.2 Hz

Figure 4: Visi-Pitch Displays Chinese Tone 4

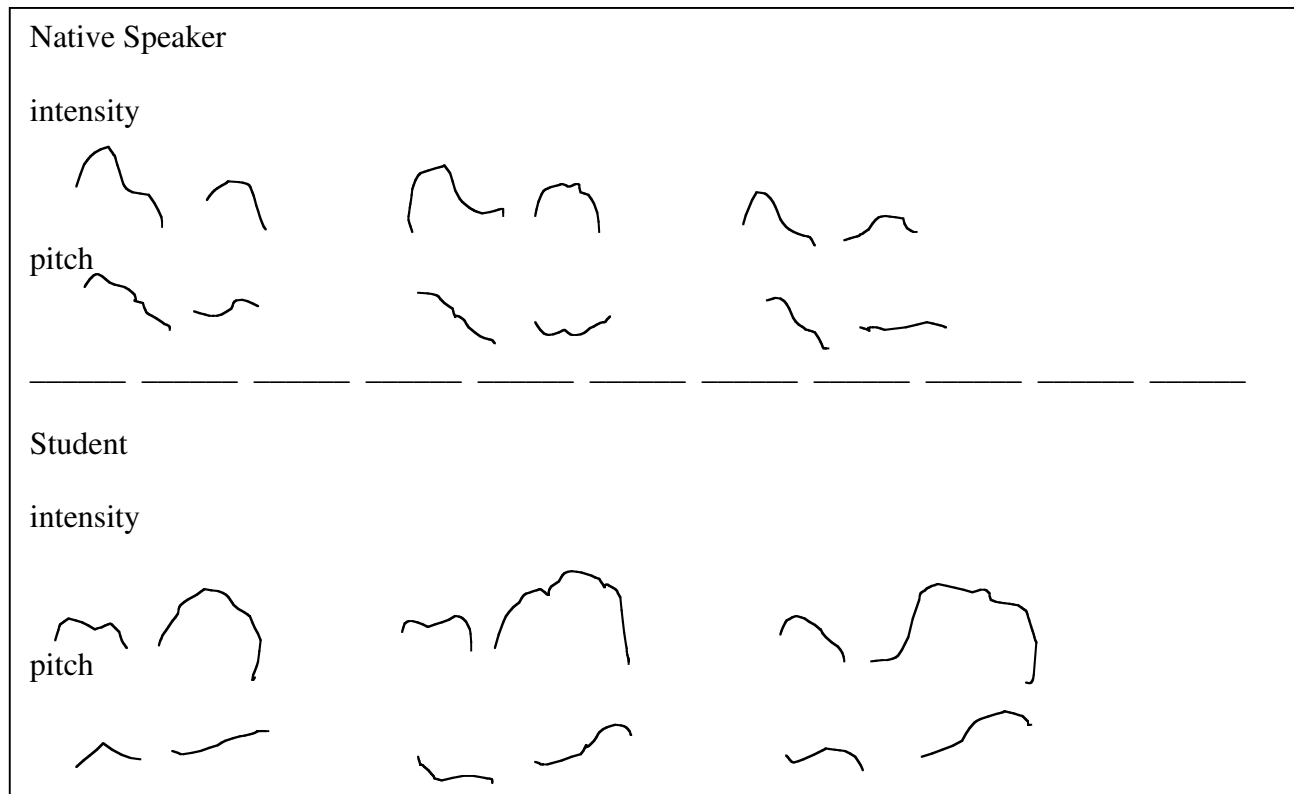


Figure 5: Visi-Pitch Displays Chinese word *wenti*

The student's attempts reflect American English pitch and intensity patterns: more stress on the second syllable and a continuously rising question intonation pattern.

Measure and display of suprasegmentals can also be useful for teaching intonation in non-tone languages such as French and German. The suprasegmental features of pitch, duration and amplitude are used to stress a word or syllable in phrases and sentences as well as to comprise overall sentence intonation. Figures 6 and 7 (from Fischer 1986b) contain Visi-Pitch examples of two questions in French as spoken by a native speaker and an American learner. For both questions, a split screen graph shows what a native French speaker said (top) and what the American learner said (bottom). For the first question *Qu'est-ce qu'il fait?* (Figure 6), only the pitch curves are shown. For the second question *Comment allez-vous?* (Figure 7), both stress and intonation patterns are shown.

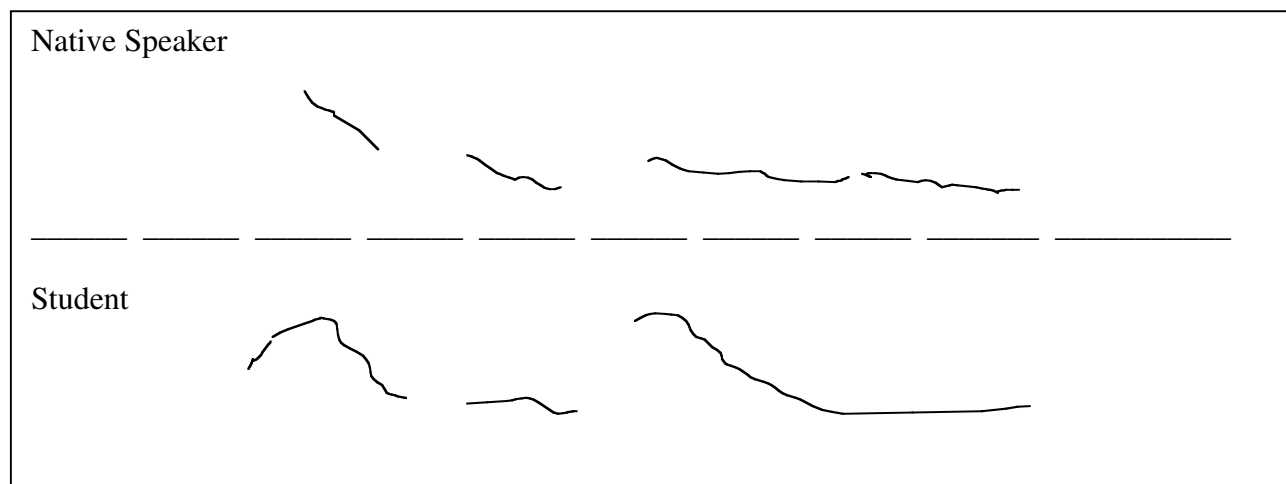


Figure 6: Visi-Pitch Displays of Pitch
French Question *Qu'est-ce qu'il fait?*

For both questions, the intonation used by the native speaker falls throughout the utterance. However, the American student asking the same questions used a more "dynamic" pattern with two peaks in pitch, as might be done in the English questions *What's he doing?* and *How are you?* The French speaker applies fairly even, slowly diminishing syllable stress, whereas the English speaker again uses a more dynamic stress pattern and applies primary stress to the first syllable of each word and less stress to the second syllable. By using visual feedback, students can be made aware of these not-so-subtle differences and can work to

improve their accents when speaking a foreign language.

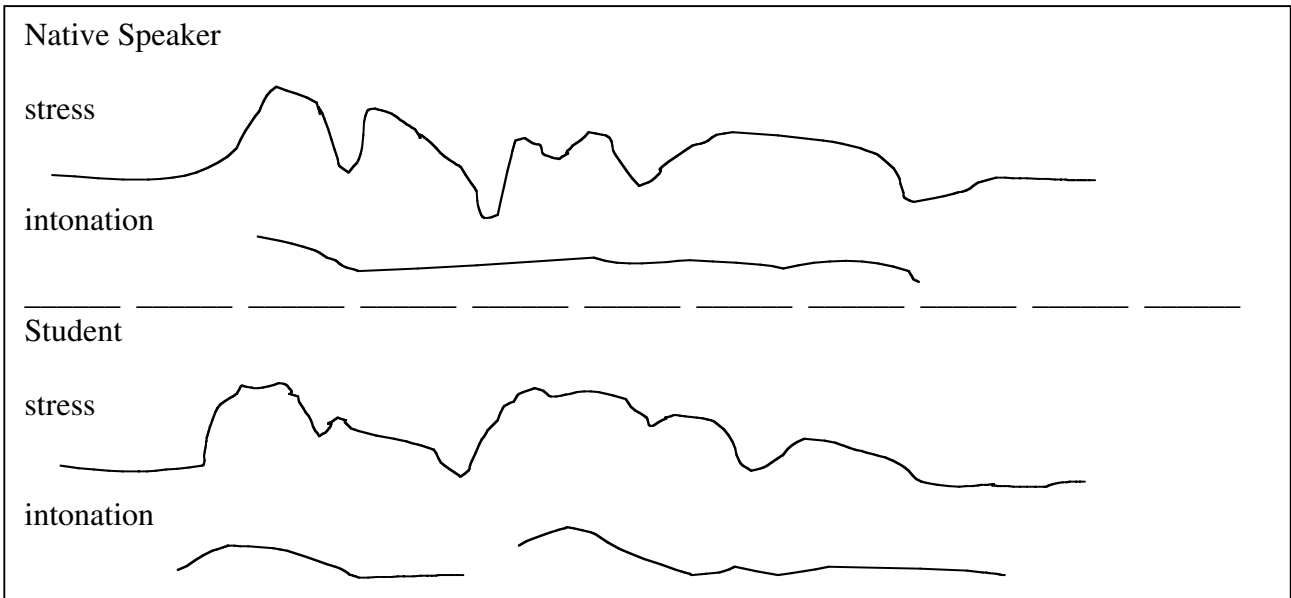


Figure 7: Visi-Pitch Displays
French question *Comment allez-vous?*

Figure 8 is an example of German sentence intonation; a short sentence was uttered by a native German speaker (top curve) and an English-speaking learner (bottom curve) using "nonsense" syllables consisting of the vowel /a/ (for greater clarity of the resulting Fo curve). The intonation used resembled that of the sentence *Auf Wiedersehen* 'good bye.' Each utterance was digitized with SoundWave and pitch-tracked with Signalyze (see below for descriptions of these programs). While the two curves are very similar, there are subtle differences, reflecting the subtle differences between English and German intonation: in German, sentence-final intonation falls more abruptly, and sentence-stress begins earlier in the stressed syllable than in English. In the top curve (German native speaker), the peak occurs sooner, and the fall is steeper,

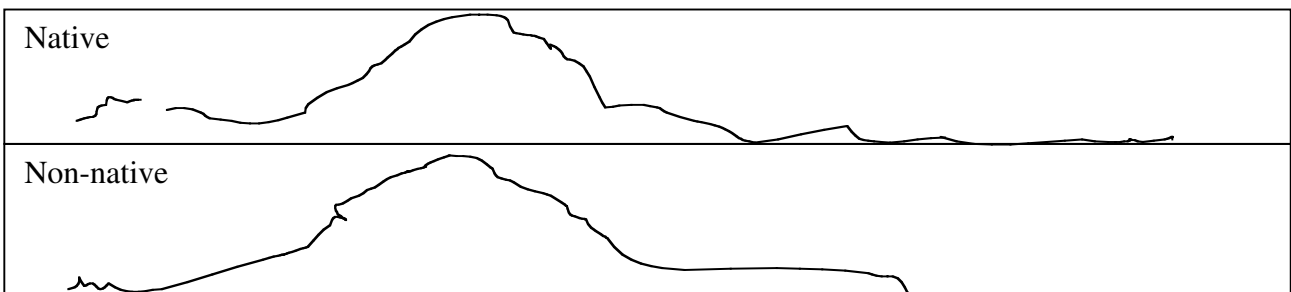


Figure 8: Pitch Curves from *Signalyze*
German nonsense utterance with /a/

less gradual, than in the bottom curve (English native speaker).

In summary, this type of instrumentation and visual feedback, though well-established in the field of Speech Pathology and Deaf Education, is relatively new to the field of foreign language instruction, and its importance has only recently been recognized.²

Hardware and Software Available for Teaching Tone and Intonation

Until recently, it was prohibitively expensive even to consider developing interactive audio-visual material for teaching tone and intonation. Much of the hardware and software was tied to costly oscillographic and spectrographic equipment or mainframes and minicomputers (e.g., Vardanian 1964, James 1976 & 1979). Today, the hardware, particularly sound digitizers, is becoming affordable, and software, in particular pitch extraction programs, is being authored for microcomputers. In this section, I summarize what is currently on the market or to appear soon on the market. (All prices quoted are as of June, 1989, except as noted.)

A. Visi-Pitch from Kay Elemetrics

Kay Elemetrics Corporation in New Jersey (address in Appendix) has developed a unique, portable, self-contained tool for teaching intonation. Visi-Pitch, model 6087PC is a stand-alone computer with a speech/voice analyzer built right in. Until now, it has been used primarily by speech pathologists and therapists, but it is gaining popularity among ESL and foreign language instructors, and other speech researchers. It extracts the fundamental frequency and amplitude features of speech signals and displays them in real-time on a built-in color monitor to provide visual biofeedback to the user. As seen above in Figures 3-6, a split screen is available for a model utterance and user input. On-screen cursors, both vertical and horizontal, facilitate precise measurement of the input signals. These signals, along with corresponding statistical calculations, can be stored and retrieved on the screen for later comparison or analysis by the student or the teacher. Hardcopy documentation can be produced with a standard dot matrix printer or video printer, and printouts of both graphics and statistics are possible. Since the model contains its own microprocessor, a word processing program and some IBM programs are also provided with the package. The prices for the computer and software are \$4,300 (without a waveform capture option) and \$4,900 (with it).

However, if one already owns an Apple 11(e or +) or an IBM PC, Visi-Pitch model 6095 or 6097 can be purchased to attach to a personal computer and costs less than the 6087PC model. For the hardware and software and an interface board, the price is \$2,950 for either the Apple II or the IBM version. When connected to either of these microcomputer hosts, programmed instructions and native speaker speech patterns can be stored on diskette.

Students can drill and practice intonation patterns and receive split-screen displays of the target or model curve on top and their repeated attempts directly below. In addition to use for tone and intonation practice, simultaneous display of pitch and amplitude show the acoustics of accent, e.g., the difference between 'desert and *des'sert* in English or 'August and *Au'gust* in German. The instrumentation works interactively with audio tape recorders, videotape, classroom monitors, and other such peripherals. In addition, the serious computer programmer can use the software routines as utility programs, can reference the stored data, and can then develop customized applications.

B. For the IBM (MSL, ILS, CSRE, CSpeech)

While Visi-Pitch was developed expressly for pedagogical purposes, it does not have the flexibility to do a wide spectrum of linguistic research. In the next two subsections, hardware and software available for IBM and Macintosh personal computers, which allow for basic acoustic phonetic research as well as for pedagogical applications, will be discussed.

One of the most popular packages for speech analysis on the IBM is "Micro Speech Lab" (MSL) for use on the IBM PC, XT, AT or compatibles, developed by the Centre for Speech Technology Research at the University of Victoria, Canada and marketed by Software Research Corporation (see Appendix for address). The system requires an EGA or VGA Graphics Card (or the Hercules) and a color monitor. The basic MSL package is U.S.\$1,600 (\$2,200 Canadian), but with all of the additional software, it totals \$3,035. MSL is a complete hardware/software system for the capture, playback and analysis of speech (and other signals). The MSL package consists of a software diskette, internally mounted 8- and 10-bit data acquisition hardware, including anti-aliasing filters, analog-to-digital (A/D) and digital-to-analog (D/A) circuitry, a microphone, and a headphone. The software includes user control of signal input, numerous waveform displays, audio output, data analysis (spectrum, pitch, and amplitude), and file management. Once an analog signal has been captured, it is converted to digital time series data which allows the signal waveform to be displayed on a graphics monitor. Editing capabilities in both graphic and audio modes allow the user to isolate portions of captured data for further analysis, and the results of all analyses are available in visual as well as numeric form. There are several optional programs: MSLEDIT extends the listening capabilities so that up to five separate files or segments of files can be displayed simultaneously, played back, cut, spliced or concatenated in any permutation. MSLPITCH displays data as waveforms and analyzes them to extract, display and store fundamental frequency (pitch) values. MSLAUDIO sets up lists of files for automatic presentation, for example, to listener-judges in

psycholinguistic experiments. MSLSPEC extends the spectrum analysis capabilities. And MSLI/O provides the source code for the data input program and a demonstration program to aid those wishing to write customized software for direct control of the hardware.

Another package on the market for the IBM is the "Interactive Laboratory System" (ILS) developed by Signal Technology in Goleta, CA. The ILS system is available for the IBM PC/XT/AT and select compatibles (as well as for the VAX, MicroVAX, VAXstation, MASSCOMP, SUN, APOLLO). Prices for the full package are \$1,875, educational price, and \$3,100, commercial price. The features included in this software system include data display and editing, digital filtering, spectral analysis, speech processing, and pattern classification. ILS also supports data acquisition, file management, data manipulation and graphics, but one should note that the package does not include hardware or software for the digitization of speech.

A review of ILS Version 6.0 for the IBM PC in the electronic phonetics journal foNETiks found it "extremely difficult to use and full of bugs, The software basically consists of routines written for mainframe computers which have been shipped 'lock-stock-and-barrel' to DOS. The use of this system defies even the most competent programmer, and the manuals do little to make things clearer"³ A demonstration diskette is available; the address of the vendor is listed in the Appendix (though this software package cannot be recommended at this time).

There are a number of other software packages, many of which are university-sponsored projects; several are currently on the market, but some will not necessarily be marketed. One system for university researchers which was released in August, 1988, is the Canadian Speech Research Environment Project (CSRE) developed at The University of Western Ontario (see Appendix for address). CSRE requires an IBM AT or compatible with EGA or better graphics (most of it will work on an XT, but very slowly; it is fastest on a 386 machine). Also required are a Data Translation 2801-A D/A, A/D board and a mouse. Included in the software are: 1) a series of three waveform editors, 2) a parametric formant synthesizer, based on KLATT, 3) three FFT-based (Fast Fourier Transform) spectral analysis programs, and 4) a basic experiment generator, to facilitate the development of programs to control experiments using speech signals. The software and documentation cost \$200 (as of August, 1988),⁴ and the DT 2801-A board sells for approximately \$1,200 (as of April, 1988).⁵

Another system, CSpeech, was developed at the University of Wisconsin, Madison (see Appendix for address) and requires an IBM compatible with a hard disk. The graphics adapters it supports include the Enhanced Color Graphics, the Hercules, and the AT&T monochrome standard. The A/D boards it supports include the DT-2801-A board, the DT-2821, the Lab Master, or the Metrabyte DAS 20. This system is available for \$1,800. While it is not inexpensive for

software, one academic user at Indiana University feels "it is the best speech analysis system available for the IBM AT... it is really quite an elegant package. it is much more manageable than ILS [which has poor wave-editing], ... is very user-friendly, ...[and] allows you to sample up to eight channels simultaneously."⁶

C. For the Macintosh: Signal Analysis "For the Rest of Us"⁷

For the Macintosh, the most widely used products on the market are MacRecorder/SoundEdit and SoundWave for digitization, and the more sophisticated MacADIOS/MacSpeechLab package for both speech acquisition and analysis. Presently under development is a new, inexpensive program for pitch extraction and display called Signalyze, which can be used in conjunction with any of the above digitizers.

MacRecorder is an 8-bit, fully integrated recording unit for the Macintosh 512/MacPlus/MacSE series developed by, Farallon Computing in Berkeley, CA (see Appendix for address).⁸ While it can be used on a 512K machine, it would be restricted to fairly short signals (e.g., approximately 15 seconds of speech if sampled at 10 kHz). Since serious research as well as any kind of courseware involves recording large quantities of signals, a hard disk would be mandatory. The package, which retails for \$199 (and for approximately \$249 for version 2.0 to appear 6/26/89), includes a recording unit that fits easily into your hand, two computer programs (MacRecorder/SoundEdit) for recording and editing, a manual, and some connectors with phono plugs. The recording unit contains a microphone, an amplifier, an antialiasing filter and the 8-bit A/D converter, all housed in the same small box. The unit needs no external power and may be connected to an outside sound source or to a high-quality microphone. The two recording programs are a Hypercard stack called MacRecorder which simulates a cassette recorder, and a more professional stand-alone program called SoundEdit, which allows for editing and storing of sound files. Advantages of this package are, first of all, its price, but also its capacity for certain types of serious acoustic research, namely measuring voice onset time (VOT), utterance, syllable, vowel or fricative durations, and narrow-band section measurements of fundamental frequency. However, it is barely sufficient and perhaps insufficient for wide-band spectrograms and formant analysis, as there is a noticeable background hiss. It would probably also be unsuitable for serious experiments in auditory perception. MacRecorder records sounds at 22 kHz but has options to record at 11, 7.3 or 5.5 kHz as well.

A similar product is the audio-digitizer with SoundWave software developed by Impulse, Inc. and marketed by AuthorWare in Minneapolis, MN (see Appendix for address), which also sells inexpensively for \$199.95. The

digitizer is also an 8-bit recording unit that generates 'snd' resources, which in turn can easily be played back through Hypercard. However, like the MacRecorder package, one can get the waveform and can calculate fundamental frequency or pitch "by hand," but actual pitch extraction is not a component of the software. To get around this, one could digitize speech files with SoundWave, then use the MacSpeechLab or Signalyze programs (discussed directly below) to perform pitch extraction, narrow and wide band sections, and narrow and wide band spectrograms. Signalyze can read SoundWave files (as well as a host of other formats--see below). Eric Keller at the University of Quebec at Montreal, the creator of Signalyze, has also written an 8-16 bit conversion routine that takes 22 kHz, 8-bit SoundWave or MacRecorder and resource files and converts them into MacADIOS/MacSpeechLab-compatible 12/16-bit files (see Appendix for Keller's address).

GW Instruments of Somerville, MA markets a wide range of hardware and software speech analysis products for the Macintosh. Their MacADIOS/MacSpeechLab packages are more sophisticated than the MacRecorder or SoundWave packages, first, because the recording system permits 12/16-bit signal acquisition, and secondly, because pitch extraction is possible. The main advantage of the 12-bit system is that spectrograms are much cleaner and formants are more reliably identified. The disadvantage of these packages is their price. For standard Macintoshes, the basic package (GWI-MSL-1, \$3,550) includes MacSpeechLab I software, MacADIOS 411 digitizer, microphone, speaker, record and play amplifiers, antialiasing filters, cables, and documentation. (For just the digitizer and software, the price is \$2,500.) For Macintosh 11 computers, the basic package (GWI-MSL-11, \$4,990) includes MacSpeechLab 11 software, MacADIOS 11 data acquisition board (\$1,490 for the board alone), MacADIOS II antialiasing filter daughterboard, microphone, speaker, record and play amplifiers, cables and documentation. MacSpeechLab permits 12/16 bit signal acquisition at 5, 10, 20 kHz, playback, time waveforms, fundamental frequency plots, wide and narrow band spectrograms, FFT spectral splice Plots, plus extensive signal editing. The Macintosh II version of MacSpeechLab in addition provides 40 or 80 kHz signal recording, amplitude and energy envelopes, and LPC spectral splice plots, Plus spectral displays at 256 gray scale levels, mappable onto excellent 300 dot-per-inch laser printouts.

The most recent analog interfaces for the Mac II are available from two California companies, Digidesign and Spectral Innovations, but both are considerably more expensive than the MacRecorder and SoundWave products. Digidesign provides for the Mac SE and Mac 11 series a Motorola fixed-point 56001-based DSP (digital signal processing) card, called Sound Accelerator. Format is 16 significant bits, and sampling rates are variable from DC to 156 kHz (stereo) or DC to 312 kHz (mono). The basic board retails for \$1,295, and the A/D

input box (converter) is \$995. Sound can be recorded and edited with a program called Sound Designer (\$349). Spectral Innovations markets for the Mac 11 series a MacDSP32 processor board (with a basic signal analysis software package) for \$2,200, and a 16-bit A/D, D/A board with antialiasing filters for \$495. These products perform calculations in the more precise floating point format and thus incorporate very fast floating point division and a number of on-board transcendental functions.⁹ For programmers, all of the libraries can be purchased for \$295.

Signalyze, a program currently in the beta-test stage to be released in September 1989, is a welcome new, multi-channel speech analysis program for the Macintosh.¹⁰ It runs on all types of Macintosh computers from the 512 series through to the IIx. Since it is an analysis program (not a data acquisition program), it does not require any special hardware, but data must be acquired with a compatible digitizer. Recorded data in a number of different formats can then be read into the program. The beta test version 0.60 of Signalyze supports the following data formats: MacRecorder (SoundEdit), SoundWave, MacADIOS (MacSpeechLab), a standard Apple and a generic sound resource format, as well as ASCII (numeric format). Planned for the commercial Version 1.0 release is the capability to read Digidesign and possibly Spectral Innovations, Audio Interchange Format (AIFF) and MacNifty files as well.

In keeping within the typical price structure for Macintosh programs, Signalyze is very affordable: it will appear in September, 1989 at a cost of \$250 (fixed), with yearly upgrades costing \$50. A prepublication price of \$150 (before September 1, 1989) is available for the current 0.60 version and for upgrades through version I.O. Signalyze is also switchable into English, French or German via a single menu command; menus, dialog boxes and error reports are all available in the chosen language.

As with all signal analysis, a fair amount of memory is required. While Signalyze can run minimally with 512K, it can take advantage of all memory available in larger machines, in that longer segments of speech can be accommodated. For example, on a MacPlus with one megabyte of internal memory, there is typically about 300K available for signal analysis. At 16 bits per sample using a 10 kHz sampling rate, this accommodates 15 seconds of speech at once. With 2.5 megs of memory, this increases to about two minutes of speech. Using a lower sampling rate can also increase the length of the sample.

Signalyze is based on the "card" principle popularized by HyperCard: there are as many channels as the user wishes to define (up to 100), but only as many channels are displayed at any one time as there is space on the screen (e.g., four on a standard Mac, ten on a 19-inch screen). Also, the display is independent of the "underlying" signal, A user can temporarily suppress a display, keeping the signal in memory. At a later point, the signal can be redisplayed. In addition, a signal can be displayed on several channels or cards

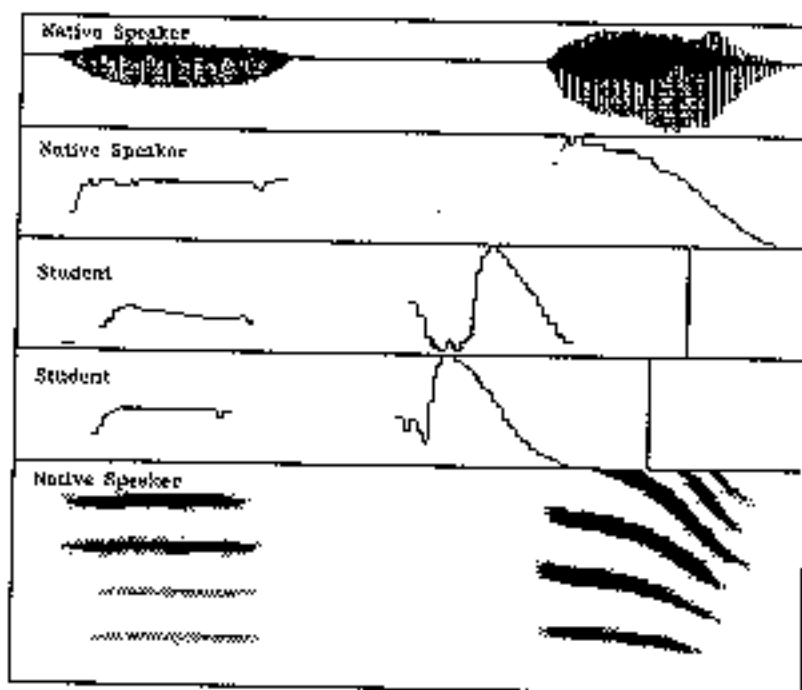
at the same time. This permits the grouping of relevant signals on a single card. Cursor movements in one display of a signal are automatically also executed in the second display of the same signal.

The program performs the following analysis functions: 1) spectral analysis - a 512-point FFT spectrogram; 2) pitch extraction—two routines are available—one based on temporal structure analysis, which is distinct from other temporal structure analyses in that it incorporates four major indicators of Fo (not just one or two), in addition to being very fast; and a second, which is the KLATT routine; 3) other signal processing routines for, e.g., amplitude envelope and zero-crossing frequencies, two types of splines, and convolutions, derivative, desampling, prosampling and limiting algorithms. It competes well with MacSpeech Lab since it offers several advantages not found in MaeSpeech Lab, e.g., extremely easy manual scoring (i.e., easily scored-and-stored numeric values under the cursor). Moreover, text/data files are integrated easily with existing word processing programs such as Word and MacWrite and with statistical packages, such as Excel, StatView, and SYSTAT. In addition, other advantages over MacSpeech Lab are that it allows the possibility of manipulating as many signals at a time as your memory permits and has faster pitch extraction as well as a plethora of conversion possibilities. It is also open to user customization, as programmers can obtain all of Signalyze as a Think C project and can add their own extensions to the program (e.g., new routines, new menus, new dialog boxes, etc.).

Figure 9. shows output from Signalyze. The top diagram is a plot (showing intensity) of the signal of a Chinese woman saying the nonsense disyllabic word *aka*. The second slot shows the Fo or pitch curve of the word, which consists of a level tone on the first syllable and a falling tone on the second syllable. The bottom slot is a narrow band spectrogram of the native speaker's model. The third and fourth slots show pitch curves of an American English speaker. In the first attempt, the first syllable was not high enough in pitch, and the second syllable fell too abruptly and was too short in duration. In the second attempt, the first syllable was a little higher, and the fall in the second syllable was a little more gradual, but in both attempts, the duration of both syllables was not nearly as long as the native speaker's. This can be seen from the statistics below the plots, which were recorded in the Signalyze program, stored as a text file, and subsequently edited with a word-processing program. (Data can also be analyzed with statistical or spreadsheet packages.)

Developing Courseware for Teaching Tone and Intonation

With the current accessibility of economical hardware and software for speech analysis, the time is ripe for developing courseware which takes advantage of these direct input and feedback capabilities. In order for



Native's duration:		DIFFERENCE:		
Vowel 1 onset	002.88	Offset	302.97	240.09 msec
Vowel 2 onset	009.75	Offset	350.36	240.61 msec
Student's duration:				
Vowel 1 onset	089.56	Offset	245.80	156.24 msec
Vowel 2 onset	476.37	Offset	690.71	214.34 msec

Figure 9: Output from *Signalvze-Chinese nonsense word aka*

courseware utilizing "pitch visualizers" to be effective in teaching or improving intonation production, certain minimum requirements must be met. This section discusses the considerations that courseware developers must bear in mind. Let us start by briefly considering the rationale for trying to perfect this aspect of pronunciation at all. For tone languages, it is obvious that correct production of the suprasegmentals is crucial for meaning; for intonation languages, the importance of appropriate and native-like speech melody has only recently begun to be realized. Even if the individual sounds and words of a language can be pronounced correctly, a foreign accent will still be evident if one simply transfers one's native intonation patterns to the foreign language (Grover et al. 1978, Tarone 1978).

More importantly, recent reexaminations of the status of pronunciation in language teaching have pointed out the need for supplementing the traditional phonemic-based view of pronunciation with a broader, discourse-based view (cf. Pennington and Richards 1986, Chun 1988). This parallels the major shift in emphasis in second and foreign language pedagogy in general from the structure and form of language to communicative or discursive meaning. A discourse-based view considers not only the individual sounds of the language, but also the prosodic features (intonation, stress, rhythm) and the voice quality features (e.g., whispery voice, high-pitched voice, falsetto, husky voice). While "all handbooks on the teaching of pronunciation agree that correct pronunciation of a foreign language (L2) cannot be achieved without complete control of the intonation, i.e. the variation of pitch, ... it is still not quite clear how intonation should be taught" (de Bot 1983:33 1). Computer courseware providing audio and visual feedback on intonation Patterns offers one viable and effective solution.

Abberton and Fourcin (1975, 158-9, who dealt with teaching rhythm and intonation to the profoundly deaf) proposed the following design requirements for a visual feedback system: 1) the visual image that is fed back has to be clear and interpretable; 2) the feedback has to be Provided in "real time" (i.e. with a minimum of delay between the production of a speech Signal and its visualization); 3) pattern proportions must be similar for speakers with different pitch ranges; 4) the information presented must not be too detailed; 5) the equipment has to be able to display a Fo (fundamental frequency) contour with the learner's imitation of it on the same screen; finally, 6) the equipment has to be inexpensive, reliable and easy to operate. All of the conditions except for the third concern the capabilities of the actual hardware and software. Condition 3 will be discussed separately below, along with other pedagogical principles that must be taken into account.

The ideal equipment currently on the market is the Visi-Pitch line of products; it meets all of the five hardware/software criteria but one: its cost can be prohibitive. Let us then consider the least expensive hardware/software combination, e.g., the SoundWave (or MacRecorder) digitizer and the Signalyze software. Condition 6 is partially met, as SoundWave (and MacRecorder) cost approximately \$200, and Signalyze will be marketed at \$250. Both SoundWave and Signalyze are easy to use, though accuracy with the pitch tracking is not yet 100% (one still does get better results with voiced segments rather than unvoiced, and sentence-final intonation with any laryngealization is not accurately displayed).¹¹ Conditions 1 and 4 are also met: if the pitch extraction and display (i.e., not the spectrograms) is used for intonation feedback, the visual feedback is reasonably clear, simple, and interpretable. The spectrograms are somewhat more complicated to read and interpret, but students will probably not need to use these. Condition 5 is also easily met, as the multiple number of display slots

allow the learner's imitation to be displayed on the same screen as the model.

The only problematic condition is the 2nd, i.e., that feedback has to be provided in "real time." Weltens and de Bot's research (1984, 87-88) showed that feedback delay (i.e., the time span between the production of an utterance and the plotting of its fundamental frequency or pitch contour) was not a critical factor, though the delays they tested were only 40 and 250 msec. Their experiment showed that "pitch visualizers do not have to be real-time in order to be effective; plotting an Fo contour with a delay of 250 msec relative to the beginning of the utterance, or even after the end of the utterance, appeared to be at least equally effective." While Visi-Pitch provides virtually immediate feedback, it costs nearly seven times as much as the SoundWave/Signalyze package, though the latter has the distinct disadvantage of longer delays in feedback, perhaps a minute or two. Using SoundWave/Signalyze, students would first have to digitize an utterance, then pitch-track it, compare it with a model, and then repeat the same process if necessary. This is time-consuming and cumbersome, but the situation should improve in time as pitch-extraction and display get faster and more accurate, and as more digitizers are supported by the pitch-analysis software.

Aside from the practical considerations of getting speech in and out, i.e., how students input speech and then receive visual feedback, there are pedagogical recommendations for the design of courseware. Condition 3, stating that pattern proportions must be similar for speakers with different pitch ranges, would require either 1) that guidelines for the amount of variance from the "norm" be set up in advance so that students can know whether their utterances "match the models well enough" or 2) that a teacher or monitor work together with students to determine whether the students' intonational contours compare satisfactorily with those of the model. In other words, students need to know the minimum amount of rise or fall in pitch and the minimum duration of any particular syllable or word which would be considered "acceptable." This is to be sure not a trivial judgment, and courseware designers must consult with native speakers and/or intonation specialists in order to determine the acceptable ranges for various parameters. Both Visi-Pitch and Signalyze can provide quantitative data on speech signals: e.g., learners can get instantaneous information on fundamental frequency (Fo) in Hz for any point or segment of the pitch curve, and they can also mark off or highlight any portion of a curve and find out its duration in msec. Courseware would thus have to have a pre-set range for minimum and maximum changes in Hz as well as for minimum and maximum durations for a given syllable or sentence, and would have to provide learners with immediate feedback as to whether or not they were within the acceptable range.¹²

Assuming the premises that intonation should be part of comprehensive language instruction and that visual feedback is an optimal pedagogical tool, the

next question is whether the criteria for courseware differ from those for teaching the more traditional aspects of language (grammar, vocabulary, listening and reading comprehension). In general, the following principles should be followed:¹³ 1) courseware should include both perception and production aspects of intonation, and tasks should proceed, preferably, from perception or comprehension to production, i.e., learners would first hear different intonation patterns, then be asked to recognize their intended meaning, make judgments about them, or predict what might follow; following that, they would be asked to imitate models and would receive visual feedback; finally, they would be given new situations and would be asked to produce appropriate responses; 2) as mentioned above, with the ever-improving technology, visual feedback should be as immediate as possible; 3) judgment of the acceptability of the input should be based on a range of values, depending on typical male vs. female Hz averages and on individual speech tempo; 4) feedback should be specific and interactive, e.g., learners should be told "You need to make the pitch rise more on syllable X" or "You need to make your pitch fall gradually over X msec for syllable Y"; 5) material should be primarily authentic, and meaningful situational or discourse context should be provided, i.e., learners need to recognize definite speech acts in dialogues and to know the exact situation in which their utterances are used or the precise attitude or emotion that is being elicited; and finally, 6) built into the courseware should be an extensive system of record-keeping so that the data can be used for research into the efficacy of the courseware, the most common difficulties of learners, and the processes by which learners master intonation.

Conclusion

While previous studies from the last thirty years have shown pitch visualizers for teaching tone and intonation to be effective, the technology to implement this pedagogical tool is only now becoming accessible and affordable to the language learning community at large. Of particular promise are the inexpensive, user-friendly hardware and software available for the Macintosh line of computers, which can already be used for a wide spectrum of acoustic phonetic research. A great advantage of computers is that they have immediate and unlimited playback capacity (as compared with cassette or reel-to-reel tapes which must be rewound for each playback and which deteriorate with repeated use). In addition, with Hypercard for the Macintosh, writing programs for foreign language CALL becomes simpler and feasible for programmers as well as for linguists and language educators. However, the development of courseware has only just begun, and there remain major hurdles to designing visual feedback systems, namely implementing digitization of speech and the immediate display of intonation curves and determining the range of "acceptable" intonational

input. In addition, built into the courseware should be a mechanism for the recording and storing of all data input so that further research on the effectiveness of computer-assisted intonation instruction can be conducted.

Appendix

Note: all prices quoted are as of June, 1989, except as noted.

VISI-PITCH FROM KAY ELEMETRICS CORP.

Kay Elemetrics Corp.
12 Maple Ave.
Pine Brook, NJ 07058-9798
Tel: (201) 227-2000 TWX 710-734-4347
FAX (201) 227-7760

Visi-Pitch 6087PC (Portable unit with built-in computer functions and color display) \$4,295 (with waveform capture option) or \$4,890 (without waveform capture option)

6095 Visi-Pitch (Software enhanced Visi-Pitch designed for, and requires microprocessor host, Apple II (e or +) for display and storage. Includes hardware and software.) \$2,950

6097 Visi-Pitch (Software enhanced Visi-Pitch designed for, and requires microprocessor host, IBM PC for display and storage. Includes hardware and software.) \$2,950

#####

HARDWARE AND SOFTWARE FOR THE IBM—PRODUCTS & VENDORS

1. Micro Speech Lab (MSL) for IBM-PC/XT/AT
Centre for Speech Technology Research
University of Victoria
Victoria, BC, Canada V8W 2Y2
Attn: Jocelyn Clayards
Tel: (604) 721-7425

Distributed by Software Research Corp.
3939 Quadra Street
Victoria, BC, Canada V8X 1J5
Attn: Annette Wright
Tel: (604) 727-3744

Prices:
MSL \$1600 for basic hardware/software package

(Optional Software)

MSLEDIT	\$370	extends MSL listening experiment capabilities
MSLSPECT	\$450	extends MSL spectrum analysis capabilities
MSLPITCH	\$320	extends MSL pitch extraction capabilities
MSLAUDIO	\$175	extends MSL audio output/input capabilities
MSLI-O	\$120	source code and demo program for data I/O

2. Interactive Laboratory System (ILS) for IBM-PC/XT/AT

Signal Technology, Inc.
5951 Encina Road
Goleta, CA 93117
Tel: 1-800-235-5787

Prices:

\$1,875 for full package, educational price, \$3,100, commercial. Demo diskette available.

3. CSpeech for IBM compatibles with hard disk

Paul Milenkovic
Dept. of Electrical Engineering
University of Wisconsin at Madison
Madison, WI 53705
Tel: (608) 262-3840

Price:

\$1,800 (software only). CSpeech supports the Data Translation DT-2801A board* and the DT-2821, the Lab Master, and the Metrabyte DAS 20; it requires Enhanced Color Graphics, the Hercules, or the AT&T standard.

4. CSRE (Canadian Speech Research Environment) for IBM-AT

Professor Donald G. Jamieson
Chairman
Dept. of Communicative Disorders
The University of Western Ontario
London, ON, N6G 1H1, Canada
USERCES0@UALTAMTS.BITNET (note: USERCES-zero, not USERCES-oh)

Price:

\$200 for software and documentation (as of August, 1988). Required: 1) IBM-AT or compatible with EGA or better graphics (it works on an XT but is very slow; it

* DT-2801A board costs approximately \$1,200. DT-707 screw terminal panel to attach tape recorder is about \$170. A TTE fix-band low pass filter is about \$120 (prices as of April, 1988).

"flies" on a 386); 2) a Data Translation 2801A D/A, A/D board*; 3) a Microsoft or compatible mouse.

#####

HARDWARE AND SOFTWARE FOR THE MACINTOSH—PRODUCTS & VENDORS

1. MacRecorder (for Macintosh 512K and better)

Farallon Computing, Inc.
2150 Kittredge St.
Berkeley, CA 94704
(415) 849-2331

Price:

\$199 retail for the digitizer and two software programs (MacRecorder & SoundEdit) for signal acquisition and editing. Version 2.0, to appear 6/26/89, will be approximately \$249. (Does not do pitch extraction.)

Also available for \$135 from:

MacConnection
14 Mill St.
Marlow, NH
1-800-622-5472 or 603-466-7711

2. SoundWave (for Macintosh 512K and better)

AuthorWare
8500 Normandale Lake Blvd., 9th floor
Minneapolis, MN 55437
(612) 921-8555

Price:

\$199.95 for digitizer and software for signal acquisition and editing. (Does not do pitch extraction.)

3. MacAdios and MacSpeech Lab

GW Instruments
35 Medford St.
Somerville, MA 02143
(617) 625-4096 FAX (617) 625-1322

For Macintosh II computers, the basic package (GWI-MSL-11, \$4,990) includes MacSpeech Lab II software, MacADIOS II Data Acquisition Board (\$1,490 for

digitizer alone), MacADIOS II Antialiasing Filter Daughterboard, Microphone, Speaker, Record and Play Amplifiers, Cables and Documentation.

For standard Macintoshes, the basic package (GWI-MSL-I, \$3,550) includes MacSpeech Lab I Software, MacADIOS 411 Digitizer, Microphone, Speaker, Record and Play Amplifiers, Antialiasing Filters, Cables, and Documentation. (For digitizer and MacSpeech Lab I software only, \$2,500.)

4. Digidesign (for Mac SE and IIs)
1360 Willow Road
Suite 101
Menlo Park, CA 94025
(415) 327-8811

Price:

\$1,295 for basic Motorola fixed-point 56001 -based DSP card ("Sound Accelerator"); \$995 for A/D input box (converter).

5. Spectral Innovations (for Mac II series)
4633 Old Ironsides Drive
Suite 450
Santa Clara, CA 95054
(408) 727-1314

Price:

\$2,200 for MacDSP32 processor board and basic signal analysis package; \$495 for 16-bit A/D, D/A and antialiasing filters; \$295 for all libraries.

6. Signalize (InfoSignal) (for Macintosh 512K and better)
Eric Keller, Ph.D.
Brainwave Software TM
231 Belair E.
ROSEMERE, Quebec, J7A 1A9 CANADA
CompuServe 76357,1213
FONETIKS@UQAM.BITNET

Price:

\$250 for license for speech signal analysis software (commercial release September 1, 1989); \$50 annual upgrade. Pre-publication price of \$150 for Version 0.60 before September 1, 1989. (Does pitch extraction, analysis and display.)

Notes

¹ This article is an expanded version of a paper/demonstration presented at the CALICO '89 Symposium, March, 1989, Colorado Springs, CO. I am grateful to Katherine Arens for help in editing the manuscript.

² For example, a program which can be used on an Apple II+ or IIe is Visible Speech, developed by Craig Dickson and Roy Snell originally for hearing-impaired children and applied to ESL and linguistics students by John Esling (see Stevens et al. 1986, 20-21). Pitch contours, shown as waveforms, are displayed in real time from tape or microphone input. The split-screen display allows the upper trace to be stored as a model while the lower trace is repeatedly redrawn in real time for student practice. These displays are useful for working on syllable stress, rhythm, and intonational patterns.

³ Katz, William. 1988. "Note to All Researchers Purchasing Speech Software for PC's." foNETiks, Vol. 1, No. 10, p.3.

⁴ A poster describing the software can be found in D.G. Jamieson and T.M. Nearey. 1988 CSRE: A Speech Research Environment. Abstracts of the Association for Research in Otolaryngology, 228. Or, for a very brief description, see Donald G. Jamieson. 1988. "First Release of CSRE: The Canadian Speech Research Environment." foNETiks, Vol. 1, No. 6, pp. 29-30.

⁵ Mary Beckman (1988. Letters to the Editor. foNETiks, Vol. 1, No. 2, p. 8.) advises that one would probably also want to spend about \$170.00 for the Data Translation screw terminal panel (DT-707) so as not to have to do too much wiring to get the board hooked up to a tape recorder. TTE markets a fix-band low pass filter that is adequate for linguistic research purposes at about \$120.00. One would again need to use a soldering iron to fit it into a switch box that allows one to use the same filter for recording and playing back, but it is not too complicated a procedure.

⁶ Ryalls, Jack. 1988. "CSpeech and the BLISS System for the IBM-AT." foNETiks, Vol. 1, No. 2, p. 4. CSpeech does not use the BLISS system, but John Mertus at Brown University is adapting the BLISS system to the AT. The BLISS system has the KLATT synthesizer and has better Fo programs. According to Richard Schwarz (1988. Letters to the Editor. foNETiks, Vol. 1, No. 2, p. 8), "Another expensive, but powerful option is the package called ASYST from MacMillan Software (approximately \$1,500 plus filters and interface). It has to be customized, but it's fairly easy to work with."

⁷ See Keller, Eric. 1988. "MacRecorder on the Macintosh: Signal Analysis "For the Rest of Us." in foNETiks, Vol. 1, No. 3, p. 17.

⁸ See review and description in foNETiks, Vol. 1, No. 2, p. 13; Vol. 1, No. 3, pp. 17-24, Vol. 1, No. 4, p. 5.

⁹ See Keller, Eric. 1989. "Questions from the Audience." foNETiks, Vol. 2, No. 3, pp. 4-5.

¹⁰ See Keller, Eric. 1989. "Signalize for the Mac." foNETiks, Vol. 2, No. 2, pp. 27-33.

¹¹ Research by Weltens and de Bot (1984, 88) had also showed that "the quality of the pitch visualizations is dependent to a considerable degree on the speech material and on (voice characteristics of) the speaker" i.e., sentences with very few unvoiced segments resulted in significantly better pitch contours than sentences which had frequently unvoiced segments. This practical consideration of keeping voiceless sounds to a minimum must be borne in mind by courseware writers.

¹² As for duration, since each person speaks at a different tempo, the minimum and maximum duration values would have to be relative to the overall tempo.

¹³ A more thorough discussion of these pedagogical principles and specific suggestions for the design of courseware (e.g., integrating visual feedback for intonation into interactive videodisc packages) can be found in Chun and Kunz (in progress).

References

- Abberton, E. and A. Fourcin. 1975. "Visual feedback and the acquisition of intonation." Foundations of Language Development, eds. E.H. Lenneberg and E. Lenneberg. 157-165. New York: Academic Press.
- Bot, K. d. 1980. "The role of feedback and feedforward in the teaching of pronunciation - an overview." System 8: 35-47.
- _____. 1983. "Visual feedback of intonation I: effectiveness and induced practice behavior." Language and Speech 26 (4): 331-350.
- Brazil, D., Coulthard, M. and Johns, C. 1980. Discourse Intonation and Language Teaching. London: Longman.

- Chun, D.M. 1988. "The neglected role of intonation in communicative competence and proficiency." *Modern Language Journal* 72 (3): 295-303.
- Fischer, L.B. 1986a. "The use of Visi-Pitch in the analysis of Chinese language suprasegmentals." Unpubl. ms. Pine Brook, NJ: Kay Elemetrics Corporation.
- _____. 1986b. "The use of audio/visual aids in the teaching and learning of French." Unpubl. ms. Pine Brook, NJ: Kay Elemetrics Corporation.
- Grover, C., D. Jamieson and M. Dobrovolsky. 1987. "Intonation in English, French, and German: perception and production." *Language and Speech* 30 (3): 277-295.
- Hubbard, P. 1988. "An Integrated Framework for CALL Courseware Evaluation." *CALICO Journal* 6 (2): 51-72.
- James, E. 1976. "The acquisition of prosodic features of speech using a speech visualizer." *IRAL* 14 (3):227-243.
- _____. 1979. "Intonation through visualization." *Current Issues in the Phonetic Sciences*, eds. H. A. P. Holien. 295-301. *Amsterdam Studies in the Theory and History of Linguistic Science, IV*, Amsterdam: John Benjamins.
- Léon, P.R. and P. Martin. 1972. "Applied Linguistics and the Teaching of Intonation." *Modern Language Journal* 56 (3): 139-144.
- Molholt, G. 1988. "Computer-assisted instruction in pronunciation for Chinese speakers of American English." *TESOL Quarterly* 22 (1): 91-111.
- Pennington, M.C. and J.C. Richards. 1986. "Pronunciation Revisited." *TESOL Quarterly* 20 (2): 207-225.
- Stevens, V., S. Spurling, D. Loritz, R. Kenner, J. Esling and M. Brennan. 1986. "New Ideas in Software Development for Linguistics and Language Learning." *CALICO Journal* 4 (1): 15-26.
- Tarone, E. 1978. "The phonology of interlanguage." *Understanding Second and Foreign Language Learning*, ed. J.C. Richards. 15-33. Rowley, MA: Newbury House.
- Vardanian, R. 1964. "Teaching English intonation through oscilloscope displays." *Language Learning* 14: 109-118.
- Weltens, B. and K.D. Bot. 1984. "Visual feedback of intonation II: Feedback delay and quality of feedback." *Language and Speech* 27 (1): 79-88.
- Wyatt, D.H. 1988. "Applying pedagogical principles to CALL courseware development." *Modern Media in Foreign Language Education: Theory and Implementation*, ed. W.F. Smith. 85-98. Lincolnwood, IL: National Textbook Co.

Author's Biodata

Dorothy M. Chun received her Ph.D. from the University of California, Berkeley and is currently Assistant Professor of German at The University of Texas at Austin. Her areas of research are German language and linguistics, suprasegmentals (particularly intonation in tonal and non-tonal languages), acoustic phonetics, discourse analysis, and the application of acoustic phonetic research to foreign language pedagogy. She has presented papers on tone and intonation in Chinese, on discourse intonation in German and Chinese, and on CALL, and has published articles in *The Modern Language Journal*, *Die Unterrichtspraxis*, and *Yearbook of the Seminar for Germanic Philology*.

Author's Address

Dorothy M. Chun
 Department of Germanic Languages
 University of Texas at Austin
 Batts Hall 216
 Austin, TX 78712-1126
 Tel.: (512)471-4422