PHONETICS IN THE NEXT TEN YEARS

Patricia A. Keating

Department of Linguistics, UCLA, Los Angeles, CA, USA

ABSTRACT

One perspective on trends in research and training in phoentics -- that of an American linguist -- is presented.

1. INTRODUCTION

Our charge in this session is to discuss the immediate future of phonetics: what will be the main topics of research, what advances will occur, how the field will change, and similar questions. When one tries to talk about the future, it is hard not to simply survey the present, and assume that basically the same will continue. At best, one finds small, recent changes, and extrapolates from these to infer large, future changes. These limitations will hold in my paper. Also, as an American and a linguist, I will have more to say about phonetics in America and phonetics as part of linguistics; I will trust to my colleagues in this session to better represent other perspectives. (See also Kuhl [2] for yet another view.) However, to broaden my horizons, I circulated a questionnaire asking for thoughts from other phoneticians. Although this paper will not consist of a systematic report on responses to that inquiry, the responses nonetheless influence my ideas here, and a list of respondents is thanked below. A second influence has been my service as phonetician/phonologist on the U.S. National Science Foundation Linguistics Program advisory panel. Like the survey, this position gives me a wider perspective on what other individual researchers take to be the most interesting and promising areas and methods of research in the immediate future.

2. CHANGES IN WHAT PEOPLE DO AND HOW THEY DO IT

2.1. Introduction

Work in phonetics is in part driven by the development and availability of new technologies. This is very much the case today, in several ways.

2.1.1. Speech Production. First, outside of industry there is a current great enthusiasm for work on speech production, especially articulator movements, due to the X-ray microbeam facility and to the various magnetometer systems, ultrasound, and other emerging possibilities. These devices are now available primarily in a few specialized labs, but magnetometers and ultrasound may come to be found in other well-equipped labs (as are the Selspot and other movement-tracking systems). The interest of these systems is that they provide data over time which is easily processed and interpreted, and thus they fit in with the study of speech dynamics. I will rely on Professor Fujimura for more extensive discussion of this topic.

2.1.2. Databases. Second, especially outside of academia, there is a growing use of large speech databases for work in acoustic phonetics and speech recognition. Virtually all industry phoneticians mention this as an important factor for the direction of future work. Such databases, as a consequence of cheaper computer memory and faster access devices, will become more common outside of industry as well, and more of them will be assembled, particularly in Europe. This development in the availability of large corpuses of speech is symbiotic with an interest in natural, connected, speech, which I will discuss below: with carefully annotated databases, analysis of large numbers of instances of any kind of acoustic phonetic event is possible. Researchers then need not limit their studies to minimal pairs or otherwise carefully controlled wordlists. However, it must be noted that the databases available today are not typically based on natural speech, nor do they span a range of styles and genres. While it may be hoped that large speech databases will in the future contribute to the study of more natural speech, much careful thought is needed to plan databases for this purpose.

2.1.3. Computers. Third, the range of acoustic analyses available outside of the major research labs is vastly improved. Ten years ago a DEC PDP-11 computer (standard in linguistics department phonetics labs) provided one user at a time access to not-always-flexible DSP, acoustic analysis, and synthesis programs. A spectrograph provided spectrograms and power spectra, again to a single user. The typical phonetics research lab would have both kinds of machines, and non-specialist departments would have just the spectrograph. Today, these two machines have merged in function, and any department or lab acquiring a "spectrograph" can have easy access to a wide range of fast acoustic analysis. But better yet, personal computers provide other, even less expensive, means for acoustic analysis, thus allowing multiple users and even portable labs. It is possible to have an IBM- or Macintosh-based system that performs most or all of the functions of the DEC machine of ten years ago. Even five years ago these were not commonly used for speech work in N. America (Keating and Anderton [1]), but now they are being acquired even by non-specialist departments. Another factor here is the availability of several commercial packages for acoustic analysis for these machines. One further change to be expected in this regard is that the availability of inexpensive processors and grey-scale printers will make computer-generated spectrograms the norm. This development will feed another development to be discussed below, the growth in acoustic phonetic description. New computers do not just make the old phonetics lab less expensive. They also make new applications available to more users. In addition to large speech databases, these include digital imaging techniques, multi-channel data collection, articulatory modelling and synthesis, especially complex models of the vocal cords and tongue, and parallel distributed processing/neural networks.

2.2. Theory-Driven Change

2.2.1. Speech Production. It seems clear that right now the most energy and enthusiasm in academic phonetics, at least, is focused on speech production studies, and this situation is likely to continue for at least a few more years. The contribution of technology to this effort was already mentioned. Another contributing factor would appear to be some waning of interest in the dominant approaches to speech perception from the 1970s. There are two theoretical driving forces as well. Some research is based in a concern for more general biological properties of movement and motor coordination (especially at Haskins Labs). Other research seeks to provide a full description of speech, including time-varying properties. There is relatively

little information at hand about time-varying properties of articulation, and this creates a need for many new studies. The one problem I see for this area of research is that it is in conflict with the goal of moving away from "laboratory speech", towards the study of more natural speech, on which more below.

The theory and techniques developed as part of a dominant research trend generally have a "trickle-down" effect. By this I mean that work on normal adults is carried over into the areas of e.g. speech development and clinical research. We can expect this to occur in the present case of speech production. Thus some responses to my questionnaire from linguists trained in speech production research mentioned these as promising areas for future work. There have always, of course, been speech production studies of children and clinical populations, given the fact that many researchers in speech production are in Speech and Hearing departments. In the future, however, such studies should be more common. These will complement the acoustically-oriented studies of the 1980s.

2.2.2. Speech Perception. The remarks above on speech production are not meant to imply that speech perception has been abandoned. One trend is to discuss speech perception in terms of more general models of the auditory system; another is to discuss it in terms of more general models of psychological category formation. It seems reasonable to expect that these two trends will be brought together in the near future, so that the physical and psychological aspects of perception will be given equal attention. A further topic of interest is the use of dynamic information in speech and, more generally, the perception of connected speech.

2.2.3. "Real" Speech. There is growing interest in the acoustic properties of more natural speech, especially in connection with work in machine and human speech recognition and text-to-speech synthesis systems. This interest is in part a reaction against work on the perception of isolated segment contrasts, and analysis of isolated acoustic characteristics of nonsense items produced under controlled laboratory conditions. It takes several paths. One is the move away from ideal and idealized "laboratory speech". This will be a big area in the next ten years --acoustic analysis and perception of casual speech, connected speech, spontaneous (vs. read) speech; women and children, regional dialects -- the full range of speech styles and voices encountered by human and nonhuman speech recognizers. We might also see more cooperation between phoneticians and social scientists as a result.

Another path is interest in lexical access --going beyond nonsense syllables and acoustic cue manipulations to get to the task that real listeners actually use speech perception for. Here we can point to a convergence of interest between, say, Ken Stevens and William Marslen-Wilson. I think that anyone who works on representational questions will come to take lexical access more into account. And this should contribute to better interaction between phoneticians and psychologists or cognitive scientists.

A third path is interest in the prosody of connected, fluent speech. It is always said that prosody is the great problem facing natural synthesis, and also that better use needs to be made of prosodic information in recognition work. At the same time, as a result of developments over the last 10 years, there are well-established linguistic theories of prosodic structure, e.g. in the "Prosodic Hierarchy" tradition ([4], [6]), just waiting to be applied to problems of phrasing and fluent speech. I see a real opportunity here that could have as one result improved academic industry cooperation.

At the acoustic end of research on prosody, an interesting new development is attention to the time-varying characteristics of the voice source as a function of prosodic structure. Source characteristics may be seen as yet another dynamic, coarticulating property of connected speech, determined in part by "inherent" segmental specifications, in part by prosody, and in part by speaker characteristics. It seems possible to me that the same trend might extend to another global property of utterances, amplitude: modeling the combined effects of respiratory energy, source, and articulation on this continuous function could contribute to more natural speech synthesis.

2.2.4. Phonology. Phonology was one of the "phonetic sciences" that the International Congresses were originally intended to bring together with core phonetics/speech science, and while some of the other "phonetic sciences" are not so actively represented here today (e.g. music, sociology), phonology has always been included in the Congress. In the early days (the 1930s) phonology meant the Prague School, and relations between phonetics and phonology at the Congresses were apparently good. Relations declined when the mantle of phonology passed to generative phonology. However, the relation of phonetics and phonology seems to be cyclic, and for the past few years we have largely returned to a period generally characterized by mutual interest and cooperation. I feel confident that this will continue for at least a while because much of the new generation of linguists goes strongly in this direction. Some of the new work by these young people is hard to categorize clearly as "phonetics" or "phonology", and I take this as a good thing, generally. (See also [3] and [5].) There are two overall research programs in this area today. One might be called the "explanation" program, because its goal is to find phonetic explanations for cross-linguistic phonological patterns -- for example, for generalizations about segment inventories. This kind of research is represented in the recent special issue of Phonetica. The other research program might be called the "grammar" program, because its goal is to establish the phonetic structure of individual utterances in a language and the general principles that underlie it. This kind of research is represented in the recent special issue of Journal of Phonetics.

2.2.5. Cross-Language Research. There has been a good deal of cross-language comparative work recently. Some of this is concerned with testing phonological theories of rules or representations, but much of it tests aspects of strictly phonetic theories, such as theories of speech production. At the same time, the current funding mechanism for speech research in Europe (Esprit projects) seems designed to encourage cross-language research, because it encourages multinational collaboration. This research is likely to be more descriptively oriented.

2.2.6. Cognitive Science. In responses to my survey, no one had a strong reaction on this topic. Indeed, phoneticians are generally not in the core of this research community, despite their strong ties to perceptual and cognitive psychology, which are in the core, and their use of computers, sometimes seen as a prerequisite to cognitive science. Why are phoneticians not active cognitive scientists? I am not sure, but perhaps our work, by its nature, is already so interdisciplinary and has ties to so many fields that we do not need to reach out for a broader forum. Perhaps in the next decade younger phoneticians will do a better job of making the connection between phonetics,

especially speech production, and cognitive science.

2.3. Other Change

2.3.1. Language Description. There is surprisingly little thorough phonetic data available on most languages. Among phoneticians, linguists have this responsibility; yet they often work only on English. However, the increased availability of good systems for acoustic analysis should change this situation somewhat, allowing descriptive linguists who are not from major phonetic research centers to carry out more instrumental phonetic studies. Also, the revival of the Journal of the IPA as an outlet to publish descriptive phonetic work will help. Thus we should see more basic data, especially acoustic phonetic data, available on many of the more "exotic" sound types. At the same time, the development of speech databases for certain major languages e.g. in Europe will result in more thorough and comprehensive acoustic phonetic descriptions of these languages. Another development is that portable computers can be used to digitize and store audio and physiological signals, and also for acoustic analysis. Such systems make it possible to use computers for instrumental phonetic fieldwork. But it's not just that more cross-language work will be done, and that perhaps we will build up more comprehensive acoustic descriptions of certain languages. It's also that this descriptive work will be more sophisticated, in that it will be tied in more directly with basic theoretical models and with analysis-by-modeling. We are beginning to see more kinds of acoustic measures being used, especially applied to less-common sounds and sound contrasts, in the way that Ken Stevens, Peter Ladefoged, and their colleagues have been doing since the 1960s. To some extent this change follows from developments in the acoustic theory of speech production, and its application to a broader range of cases.

It is also possible that we will see changes in which languages are studied by phoneticians. At least in the U.S., languages of the Pacific, especially SE Asian languages, are taking on more importance, yet these are not much studied by phoneticians. At the other extreme, there seems to be more sensitivity among linguists generally to the fact that many indigenous languages are in or near their last generation of fluent speakers. Perhaps phoneticians will play a larger role in preserving a record of such languages.

2.3.2. Non-Phonetician Users. Another change I see happening now is that there are more of what I will call "users" of phonetic data. By "users" I mean people who learn how to use certain equipment and make certain measurements without having a more general interest in phonetics, without first having general training in phonetics, and with no long-term interest in setting up, maintaining, or perhaps even being affiliated with, a phonetics lab. The use of speech data is much more common among non-phoneticians now than before, I think. First, psycholinguists who do not study speech per se are increasingly interested in working in an auditory modality. This means they use speech signals to create their experimental stimuli. This is a striking aspect of the speech lab set-up at the Max-Planck-Institute for Psycholinguistics, for example. Second, phonologists and linguists in language-area specialties seem ready to rely more on instrumental data at times where the ear alone, especially the untrained ear, is not reliable. For example, it is once again becoming the norm to see f0 traces in essentially phonological studies of intonation or of tone. (Use of instrumental data by phonologists is generally cyclic. A facilitating factor these days is the familiarity of personal computers to non-phoneticians, which makes

the phonetics lab less threatening than it once might have seemed.) Third, second language researchers seem to be increasingly interested in acoustic analysis as a tool for contrastive analysis and for assessment. (This also will contribute to language description and to cross-language work.) As a result, researchers in all of these areas will want access to phonetics instrumentation. My impression of this situation is that laboratory phonetics is maturing as a "service" discipline, exactly parallel to the way that traditional impressionistic phonetics has become a service discipline. (I say this despite the fact that phonetic transcription is a theoretical construct, as Ladefoged has recently made great efforts to stress.) The use of phonetics lab equipment for speech data analysis or manipulation is a service that phoneticians offer to other phonetics "consumers". The good side of this is that more people get involved in phonetics, bringing more data and new issues to the phonetician's attention. The bad side of this is that non-phoneticians will come to think that this is the whole of instrumental phonetics. They may thing that, just like impressionistic auditory phonetics, instrumental phonetics is an essentially non-theoretical discipline where everything has already been worked out, and where the only remaining challenges are descriptive. They may also think that qualifications in phonetics consist mainly of technical training. Such an impression of phonetics among non-phoneticians would not be good for our field. For example, the best students would tend not to choose to go into phonetics.

2.3.3. Historical Phonetics. My own impression is that there is renewed interest among younger phoneticians in the area of sound change and historical phonetics, but not yet any unifying themes or strong theoretical innovations in their work. Generally this interest springs from work in particular language families or areas. Not only might we hope for increased phonetic sophistication in theories of sound change, but data from sound change might be expected to provide test data for general phonetic theories.

2.3.4. Funding. It seems clear that phoneticians able and willing to work with engineers have a much better time in the funding situation. The same obviously applies to phoneticians interested in more applied kinds of work. Our European colleagues benefit financially from collaborations with industry, but caution us about the problems that arise (lack of freedom to do basic research, and bureaucratic/administrative problems). Many (American) academic respondents to my questionnaire already complain that speech technology seems to be driving the field more than they would like. Almost everyone in academia responding to my questionnaire thought that funding considerations were going to become more important in the future, that is, that they would have to choose their topics based more on considerations of funding possibilities than they have in the past.

3. CHANGES IN EMPLOYMENT, ETC.

3.1. Retirements, Hiring Patterns

Although all the new reports assure us that academic jobs will be booming in this decade, and although there is more hiring going on in U.S. universities, a disturbing pattern is that there has been no increase in long-term jobs for new phonetics Ph.D.s in linguistics. Linguists are going to Speech and Hearing departments (which seems to be a genuine growth area for linguists just now), and to industry. The situation in Europe, Australia, etc. is also not especially encouraging, and essentially jobs are scarce everywhere. Perhaps we faculty

could help our students by doing more -- or leaning on these very students to do more! -- to integrate phonetics better into cognitive science, since that does seem to be an academic growth area. Another strategem is to convince phonologists in more departments that they would benefit from having a phonetician as a colleague, while also pointing out that a reasonable phonetics lab can be started for less money than might have been expected. The other thing we can do is to call more attention to the availability of absolutely top-rate women and minority phoneticians; a university can do well in its affirmative action programs by expanding in phonetics.

3.2. What Students Should Be Studying For the Future

In my questionnaire I asked respondents to advise current students on how to prepare for their futures, including what they themselves look for in hiring. The responses were quite revealing, and can be summarized as follows. Most important for phoneticians generally is a solid ground in acoustic theory and speech acoustics. Also important are basic knowledge of experimental design and statistics and user skills in acoustic analysis and synthesis. The more computer experience a student has, the better, even in academia; programming in C and in an AI language are recommended for industry. Some background in linguistics is helpful for industry, and substantial background in phonology is helpful for academic lingustics departments. Another recommendation is experience with modeling and simulation; it reinforces knowledge of theory and demonstrates computer skills. Finally, students interested in an industry career should try to arrange an internship or some other kind of student job at a company or industry-related research center. My hope is that students can use this report as a guide in considering possible directions for their own research or future careers. There are many exciting research questions to choose from, most of which can be directed to either academic or industry careers. The classic advice, to find an interdisciplinary niche, remains relevant today for both academia and industry. Students should also think about how theoretical developments from various areas might translate into applications, or contribute to progress in applications-oriented research. Students should not wait for industry to realize that their skills are needed, because they themselves can help create the market for those skills.

4. References

[1] KEATING, P. and ANDERTON, A. (1987), "Survey of phonetics computers in North American laboratories", UCLA Working Papers in Phonetics, 66, 104-123.

[2] KUHL, P. (1990), "Hot topics in speech communication", J. Acoust. Soc. Am. 88, Suppl. 1, S69.

[3] LIBERMAN, M. (1983), "In favor of some uncommon approaches to the study of speech", The Production of Speech, ed. P. MacNeilage, New York: Springer-Verlag.

[4] NESPOR, M. and VOGEL, I. (1986), Prosodic Phonology, Dordrecht: Foris Publications.

[5] OHALA, J. (1990), "There is no interface between phonology and phonetics: a personal view", J. Phonetics, 18, 153-272.

[6] SELKIRK, E. (1980), "Prosodic domains in phonology: Sanskrit revisited", Juncture, ed. M. Aronoff and M.-L Kean, Saratoga: Anma Libri.

5. ACKNOWLEDGMENTS

Thanks to those who responded to my survey: S. Baum, M. Beckman, P. Beddor, S. Boyce, C. Browman, G. Bruce, F. Chen, K. Denning, F. Ganon, J. Gilbert, S. Hawkins, C. Henton, M. Huffman, M. Jackson, H. Javkin, K. Johnson, R. Krakow, P. Ladefoged, I. Maddieson, J. Ohala, B. Poser, P. Price, B. Repp, C.-Y Tseng, M. Withgott and a couple of unidentified respondents. Thanks also to Louis Goldstein and to members of the UCLA Phonetics Lab.