

The Role of Phonetics in Linguistics and Other Fields

Arthur S. Abramson

Introduction

Phonetics is the scientific study of the sounds of speech. The person doing this work is called a phonetician or, quite often, a speech scientist. It is not surprising that some of us phoneticians are based in linguistics because of our great interest in speech as the major way in which language encodes messages. A longstanding approach among linguists doing field work can be called practical phonetics or articulatory-auditory phonetics. Some of us who work in the laboratory concentrate on the motor control of the articulators and larynx that make the sounds of speech. This needs the collaboration of physiologists or physicians. Others give special attention to the acoustic analysis of speech, which is helped by the cooperation of physicists or engineers. Then there are those who are interested mainly in the perception of speech; their work draws upon the techniques and theories of experimental psychologists. Thus we see that phonetics is an interdisciplinary field of research. All the foregoing specialties are divisions of experimental phonetics, which involves the use of instruments and computer programs to identify the elements of speech that carry linguistic information and, perhaps, information about the speaker. The phonetician might then design experiments to test the perceptual relevance of the findings.

Practical Phonetics

A longstanding tradition in phonetics has been the training of students, through hours and hours of drill in hearing and mimicking speech sounds, to reach the goal of being able to hear, produce, and write down (transcribe) all the sounds of speech possible for human beings. Those who come closest to this ambitious goal are linguists with such phonetic training who do a great deal of fieldwork on a variety of languages. This skillful use of ears, eyes, and vocal tract must be supported by a fairly good knowledge of the anatomy

of the vocal tract. When not working with a teacher, the best book for self-help is no doubt the recent one by J.C. Catford (2001). He leads the student step by step through the production of a wide gamut of speech sounds. For phonetic transcription the widely used alphabet of the International Phonetic Association (1999) is preferred, although others exist. Occasional additions to its International Phonetic Alphabet (IPA) are published in the *Journal of the International Phonetic Association*.

A problem facing the would-be practical phonetician today is that very few academic departments of linguistics offer the necessary training. Rather, the emphasis is on phonology, the study of principles leading to rather abstract sound systems of languages. Thus, the field linguist will use the phonetic details obtained from native speakers of a language as raw material for testing hypotheses on the system of abstract units that serve best to differentiate morphemes and words. A word of warning: No matter how elegantly logical, formal, systematic, and efficient the linguist's phonology as a component of the grammar of a language may be, if it fails to yield at its output phonetically appropriate utterances, it does not satisfactorily account for the sound system of that language. That is, phonetics is not merely a tool for the linguist; it is a discipline in its own right with a vital function in the study of language.

A further cautionary note is needed here on the limits to the reliability of a phonetic analysis yielded by purely practical (auditory-articulatory) methods. No matter how keen an ear and nimble a tongue the practical phonetician may have, his or her observations are constrained by psychophysical limits to perception. In addition, in both perception and production of speech, he may not have fully overcome a lifetime of native-language biases in the neuro-muscular control of articulation and the perception of linguistically distinctive differences. Thus, I must argue that while everyone working in the phonetic sciences ought to have at least a good background in practical phonetics, it is important to validate such data with instrumental studies in the laboratory and perceptual experiments, as done in the field of experimental phonetics.

Experimental Phonetics

An experimental approach to a question in phonetics ought to require a theoretical model from which hypotheses can be formed for testing. Conceivably such a model could be taken from phonology; however, we

must not simply accept the linguist's phonological categories and seek to validate them in the laboratory. Rather, the model must include a mechanism whereby the categories in question are distinguished. Such testable hypotheses often come to us in phonetics from acoustics, physiology, and psychology, as well as phonetics itself. For a short article for newcomers on the unfolding of linguistics, it seems to me that the best approach is to give a brief review of my involvement in a series of many studies on the production and perception of phonologically distinctive voicing in consonants.

Voice timing

In the early 1960s the late Leigh Lisker and I began talking about some of the confusion in the linguistic literature about the phonetic bases of such widely used concepts as voicing, aspiration, and tenseness in consonants. Voicing seemed to be defined very narrowly as solely the presence of vibrations of the vocal folds (glottal pulses) during the closure or constriction of a consonant. Thus, a consonant was said to be voiced if audible glottal pulses were present in the "segment" and voiceless if not. Aspiration was taken to be breathy noise coming through the open glottis during, typically, the release phase of a voiceless stop consonant. Tenseness was variously and vaguely defined as extra contraction of muscles in the upper vocal tract: the larynx, pharynx, oral cavity, tongue, or lips.

We reasoned that most of these confusions would go away if we took the notion of phonetic segment much less seriously. By that time, extensive research at Haskins Laboratories and a few other institutions had made it clear that the information-bearing elements in the sound waves of speech overlap a good deal in a kind of parallel transmission and are not limited to segmental time slots occupied by separate phonemes. This thinking led to the notion of the timing of the onset or offset of voicing relative to some critical point in the articulation of the consonant.

In our initial publications on the subject (Lisker & Abramson, 1964; Abramson & Lisker, 1965), we labeled the dimension voice onset time (VOT), because in a wide variety of languages the dimension was most richly exploited in word-initial position. Although VOT is now the label most commonly used, I like to point out that it is really one subtype of the larger phenomenon of voice timing (Abramson, 1977).

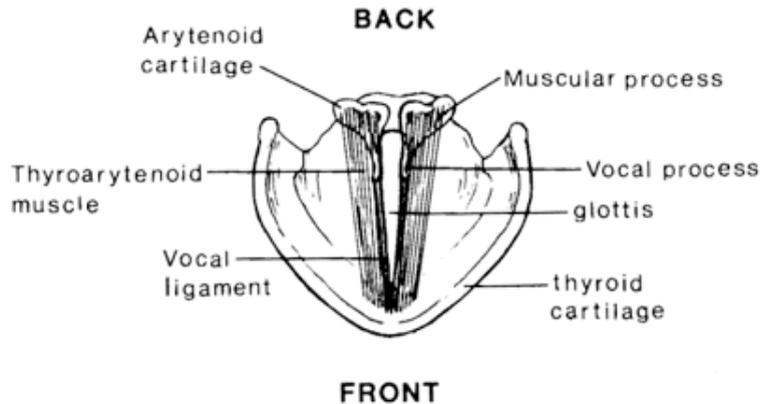


Figure 1. View of the larynx from above at the level of the vocal folds.

Let me pause here to provide some background for my discussion of VOT. In Figure 1 we are looking down at a somewhat simplified view of a cross-section of the larynx, which sits on top of the trachea (wind pipe). The two dark bands running from the inner surface of the thyroid cartilage in front to the arytenoid cartilages in back are the vocal folds. (They are also called, less appropriately, vocal cords.) When the folds are somewhat abducted (spread), as here, we can see the space between them, which is called the glottis. When the folds are adducted (brought together) in a suitable state of tension, airflow from the lungs below can set them vibrating; this is phonation (voice). Voicing can continue only as long as the air pressure below the glottis is greater than that above. If the folds are tightly adducted, we have a glottal stop with no sound until the closure is released. If the vocal folds are abducted and eddies of turbulent air come through the glottis, we have voiceless aspiration. If the front portions of the folds are adducted enough for voicing but there is an opening in the glottis toward the back, we will have voiced aspiration (murmur, breathy voice). In English, for example, the word *hair* is phonetically the same as *air*, except that its vowel has a breathy onset. Thus there are as many varieties of English /h/ as there are vowels that go with it.

With the help of Figure 2, we can now go back to the discussion of VOT.

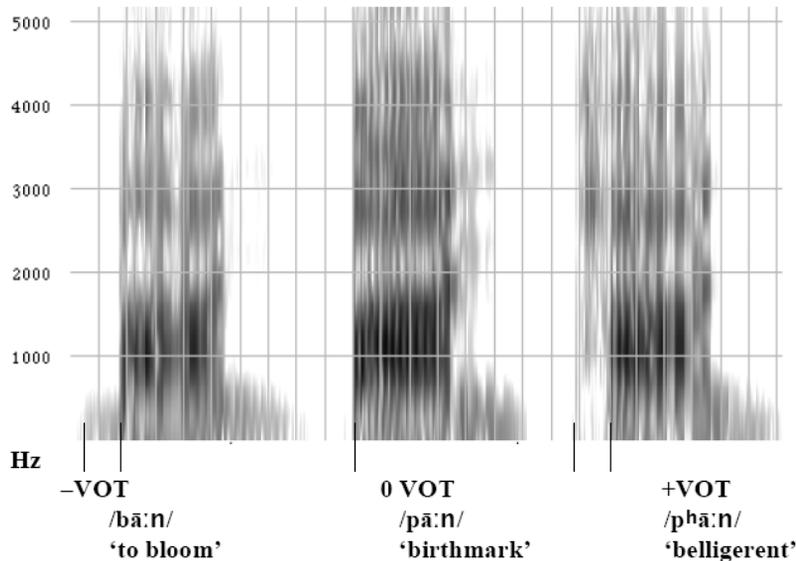


Figure 2. Wide-band spectrogram of Thai words showing negative, zero, and positive voice onset time in the initial stops.

The three Thai words in Figure 2, all on the mid tone, were recorded by a woman, a native speaker, to represent the three-way voicing distinction traditionally labeled, from left to right, voiced, voiceless unaspirated, and voiceless aspirated. The sound spectrogram shows an analysis in terms of frequency in Herz on the vertical scale from 0 to 5000Hz, time along the horizontal axis, and approximate relative intensity as degrees of darkness. Actual values for the latter two are available in the computer-generated underlying analysis. The horizontal black bands are formants, that is, resonances of the of the upper vocal tract as shaped by articulatory dynamics.

The articulatory point of reference for VOT chosen by Lisker and me is the release of the stop consonant, which can be seen in the second vertical line under the first word, the single vertical line under the second word, and the first vertical line under the third word. For voice inset at the moment of release, as in the second word, we established the conventional value of 0 ms (milliseconds). If in a given speech sample VOT occurs before the release, it is known as voicing lead and has a negative value. In this utterance of /bā:n/ VOT = -123ms. For /pā:n/ VOT = 0ms, and for /phā:n/ VOT = +108ms. Since

we are dealing here with stop consonants, it follows that the voicing lead in the first word must be occurring during the articulatory closure. In the second word, however, the closure is silent. In the third, voicing is delayed after release, and the lag is filled with aspiration noise coming through the glottis that has not yet closed for phonation.

We have demonstrated in the aforementioned sources and others (e.g., Lisker & Abramson, 1971; Abramson & Lisker, 1973) as have many other researchers (e.g., Raphael, Tobin, Faber, Most, Kollia, & Milstein, 1995; Koenig, Mencl, & Lucero, 2005) that VOT is a powerful mechanism for handling two-way and three-way voicing distinctions in a large number of languages. This is not to say, however, that VOT is all powerful in succeeding to account for every single consonantal category involving laryngeal control of some kind. There are cases that require the invoking of one or more other dimensions that may intersect with VOT.

Intersection of voice onset time with phonation type

An example of the need to supplement VOT with another dimension is found in the Indo-Iranian languages, such as Hindi. Waveforms of two Hindi words appear in Figure 3. I am using waveforms rather than spectrograms here, because in reduced size on the printed page waveforms show more clearly the points to be made. These waveforms are records of the complex variations in sound pressure of two utterances recorded by a native speaker of Hindi. Once again, time runs from left to right along the horizontal axis.

Hindi has four phonologically distinct stop categories that are traditionally called voiced unaspirated, voiced aspirated, voiceless unaspirated, and voiceless aspirated. Note that the first, third, and fourth are much the same phonetically as the three categories of Thai. Not surprisingly, we found (Lisker & Abramson, 1964) that VOT fails to distinguish the two voiced categories of Hindi. On the VOT dimension the unaspirated /bdg/ and the aspirated / b^h d^h g^h/ overlap pair by pair.

The point of Figure 3 is to show that there are two kinds of aspiration in Hindi. In the word /p^hal/ to the right we see a span of 124ms of voiceless aspiration (turbulence), as shown by the two short vertical lines on the axis at the bottom.. This is like the Thai speaker's voiceless aspirate in Figure 2. To the left in Figure 3 the first two vertical lines under the word /b^hal/ mark off a span of negative VOT (voicing lead); that is, VOT = -109ms. Note that the amplitude of the voicing lead, as shown by the vertical breadth of the trace, becomes very low just before the stop release at the second short vertical line.

relevant factors are for the differentiation of the two types in speech production.

Conclusion

Phonetics is an interdisciplinary field. Every four years at the International Congresses of Phonetic Sciences we see researchers based not only in phonetics itself but also in linguistics, speech and hearing science, psychology, communications engineering, computer science, and medicine. Phonetic research benefits from the expertise, knowledge, and techniques of all these disciplines. The discussions here of practical phonetics and one topic in experimental phonetics are meant to give some idea of the field.

References

- Abramson, Arthur S., & Lisker, Leigh. 1965. Voice onset time in stop consonants: Acoustic analysis and synthesis. *Rapports du 5e Congrès International d'Acoustique*. 1a. Conf. A51.
- Abramson, Arthur S., & Lisker, L. 1973. Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, 1, 1-8.
- Abramson, Arthur S. 1977. Laryngeal timing in consonant distinctions. *Phonetica*, 34, 295-303
- Catford, J.C. 2001. *A practical introduction to phonetics*, 2nd ed. Oxford: Oxford University Press.
- International Phonetic Association. 1999. *Handbook of the International Phonetic Association*. Cambridge, U.K.: Cambridge University Press.
- Koenig, Laura L., Mencl, W. Einar, & Lucero, Jorge, C., 2005. Multidimensional analyses of voicing offsets and onsets in female speakers. *Journal of the Acoustical Society of America*, 118, 2535–2550.
- Lisker, Leigh, & Abramson, Arthur S. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Lisker, L. and Abramson, A. S. 1971. Distinctive features and laryngeal control. *Language*, 47, 767-785.
- Raphael, Lawrence J., Tobin, Yishai, Faber, Alice, Most, Tova, Kollia, H. Betty, & Milstein, Doron. 1995. Intermediate values of voice onset time. *Production of speech: contemporary issues for Katherine Safford Harris*, ed. by Fredericka Bell-Berti and Lawrence J. Raphael, 117–127. New York: American Institute of Physics.