



Acoustic phonetics

Acoustic phonetics is a subfield of phonetics, which deals with acoustic aspects of speech sounds. Acoustic phonetics investigates time domain features such as the mean squared amplitude of a waveform, its duration, its fundamental frequency, or frequency domain features such as the frequency spectrum, or even combined spectrotemporal features and the relationship of these properties to other branches of phonetics (e.g. articulatory or auditory phonetics), and to abstract linguistic concepts such as phonemes, phrases, or utterances.

The study of acoustic phonetics was greatly enhanced in the late 19th century by the invention of the Edison phonograph. The phonograph allowed the speech signal to be recorded and then later processed and analyzed. By replaying the same speech signal from the phonograph several times, filtering it each time with a different band-pass filter, a spectrogram of the speech utterance could be built up. A series of papers by Ludimar Hermann published in Pflügers Archiv in the last two decades of the 19th century investigated the spectral properties of vowels and consonants using the Edison phonograph, and it was in these papers that the term *formant* was first introduced. Hermann also played back vowel recordings made with the Edison phonograph at different speeds to distinguish between Willis' and Wheatstone's theories of vowel production.

Further advances in acoustic phonetics were made possible by the development of the telephone industry. (Incidentally, Alexander Graham Bell's father, Alexander Melville Bell, was a phonetician.) During World War II, work at the Bell Telephone Laboratories (which invented the spectrograph) greatly facilitated the systematic study of the spectral properties of periodic and aperiodic speech sounds, vocal tract resonances and vowel formants, voice quality, prosody, etc.

Integrated linear prediction residuals (ILPR) was an effective feature proposed by T V Ananthapadmanabha in 1995, which closely approximates the voice source signal.^[1] This proved to be very effective in accurate estimation of the epochs or the glottal closure instant.^[2] A G Ramakrishnan et al. showed in 2015 that the discrete cosine transform coefficients of the ILPR contains speaker information that supplements the mel frequency cepstral coefficients.^[3] Plosion index is another scalar, time-domain feature that was introduced by T V Ananthapadmanabha et al. for characterizing the closure-burst transition of stop consonants.^[4]

On a theoretical level, speech acoustics can be modeled in a way analogous to electrical circuits. Lord Rayleigh was among the first to recognize that the new electric theory could be used in acoustics, but it was not until 1941 that the circuit model was effectively used, in a book by Chiba and Kajiyama called "The Vowel: Its Nature and Structure". (This book by Japanese authors working in Japan was published in English at the height of World War II.) In 1952, Roman Jakobson, Gunnar Fant, and Morris Halle wrote "Preliminaries to Speech Analysis", a seminal work tying acoustic phonetics and phonological theory together. This little book was followed in 1960 by Fant "Acoustic Theory of Speech Production", which has remained the major theoretical foundation for speech acoustic research in both the academy and industry. (Fant was himself very involved in the telephone industry.) Other important framers of the field include Kenneth N. Stevens who wrote "Acoustic Phonetics", Osamu Fujimura, and Peter Ladefoged.

See also

- [List of phonetics topics](#)
- [Human voice](#)

Bibliography

- Clark, John; & Yallop, Colin. (1995). *An introduction to phonetics and phonology* (2nd ed.). Oxford: Blackwell. ISBN 0-631-19452-5.
- Johnson, Keith (2003). *Acoustic and Auditory Phonetics (Illustrated)*. 2nd edition by Blackwell Publishing Ltd. ISBN 1-4051-0122-9 (hardback: alkaline paper); ISBN 1-4051-0123-7 (paperback: alkaline paper).
- Ladefoged, Peter (1996). *Elements of Acoustic Phonetics* (2nd ed.). The University of Chicago Press, Ltd. London. ISBN 0-226-46763-5 (cloth); ISBN 0-226-46764-3 (paper).
- Fant, Gunnar. (1960). *Acoustic theory of speech production, with calculations based on X-ray studies of Russian articulations*. Description and analysis of contemporary standard Russian (No. 2). s'Gravenhage: Mouton. (2nd ed. published in 1970).
- Hardcastle, William J.; & Laver, John (Eds.). (1997). *The handbook of phonetic sciences*. Oxford: Blackwell Publishers. ISBN 0-631-18848-7.
- Hermann, L. (1890) "Phonophotographische Untersuchungen". Pflüger's Archiv. f. d. ges Physiol. LXXIV.
- Jakobson, Roman; Fant, Gunnar; & Halle, Morris. (1952). *Preliminaries to speech analysis: The distinctive features and their correlates*. MIT acoustics laboratory technical report (No. 13). Cambridge, MA: MIT.
- Flanagan, James L. (1972). *Speech analysis, synthesis, and perception* (2nd ed.). Berlin: Springer-Verlag. ISBN 0-387-05561-4.
- Kent, Raymond D.; & Read, Charles. (1992). *The acoustic analysis of speech*. San Diego: Singular Publishing Group. ISBN 1-879105-43-8.
- Pisoni, David B.; & Remez, Robert E. (Eds.). (2004). *The handbook of speech perception*. Oxford: Blackwell. ISBN 0-631-22927-2.
- Stevens, Kenneth N. (2000). *Acoustic Phonetics*. Current Studies in Linguistics (No. 30). Cambridge, MA: MIT. ISBN 0-262-69250-3.
- Stevens, Kenneth N. (2002). "Toward a model for lexical access based on acoustic landmarks and distinctive features". *The Journal of the Acoustical Society of America*. **111** (4): 1872–1891. Bibcode:2002ASAJ..111.1872S (<https://ui.adsabs.harvard.edu/abs/2002ASAJ..111.1872S>). doi:10.1121/1.1458026 (<https://doi.org/10.1121/1.1458026>). PMID 12002871 (<https://pubmed.ncbi.nlm.nih.gov/12002871>). S2CID 1811670 (<https://api.semanticscholar.org/CorpusID:1811670>).

References

1. T. V. Ananthapadmanabha, "Acoustic factors determining perceived voice quality", in *Vocal fold Physiology - Voice quality control*, O.Fujimura and M. Hirano, Eds. San Diego, Cal.: Singualr publishing group, 1995, ch. 7, pp. 113–126.
2. A. P. Prathosh, T. V. Ananthapadmanabha, and A. G. Ramakrishnan, "Epoch extraction based on integrated linear prediction residual using plosion index", *IEEE Transactions on Audio, Speech, and Language Processing*, 2013, Vol. 21, Iss. 12, pp. 2471-2480.

3. A G Ramakrishnan, B Abhiram and S R Mahadeva Prasanna, "Voice source characterization using pitch synchronous discrete cosine transform for speaker identification", Journal of the Acoustical Society of America Express Letters, Vol. 137(), pp., 2015.
4. T V Ananthapadmanabha, A P Prathosh, A G Ramakrishnan, "Detection of the closure-burst transitions of stops and affricates in continuous speech using the plosion index", Journal of the Acoustical Society of America, Vol. 137, 2015.

External links

- [Speech Analysis Tutorial \(https://web.archive.org/web/20071126114817/http://www.ling.lu.se/research/speechtutorial/tutorial.html\)](https://web.archive.org/web/20071126114817/http://www.ling.lu.se/research/speechtutorial/tutorial.html)
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Acoustic_phonetics&oldid=1173326626"