Mapping the perceptual vowel space – Articulatory, acoustic and perceptual relations between vowel categories

Nadja Kerschhofer-Puhalo
University of Vienna
nadja.kerschhofer@univie.ac.at

The idea of "the human vowel space" is commonly associated with illustrations where vowel sounds are located as points in a two-dimensional representation based on the first two formants. F1xF2-based representations have been claimed to correspond to a certain degree to articulatory constellations in the vocal tract, mainly to the position of the tongue body. Feature-based definitions of vowels referring to the three major features vowel height, frontness/backness and rounding as well as the IPA quadrilateral are also closely related to the idea that the human vowel space can be visualized in a two-dimensional representation. However, it is well known that a representation in terms of points in an F1xF2-plot is insufficient and too static with respect to a number of acoustic characteristics of the speech signal such as higher formant frequencies, duration, pitch and dynamic patterns of formant movements (e.g. Hillenbrand 2013). Moreover, an equation of articulatory and acoustic dimensions is quite problematic, specifically with respect to the concept of "tongue height" (see Wood 1975, 1979; Ladefoged 1993[3]: 220f). Yet, despite these drawbacks, a two-dimensional F1xF2 representation of the acoustic vowel space is more or less standard to describe and compare vowel categories in a wide variety of languages.

In this contribution, the phonetic vowel space will be described as multi-dimensional. Multi-dimensionality refers to articulatory, acoustic and perceptual dimensions. These dimensions have no linear correspondences to each other. They do not necessarily correspond to phonological features such as vowel height, frontness/backness or rounding and may not directly be associated with specific acoustic properties such as F1 and F2, as no linear correlation of variation in phonetic parameters and perceived phonemic identity can be assumed. Therefore, it will be argued here for a careful differentiation of (a) the *articulatory space* (vocal tract), (b) the *acoustic space* and (c) the language-specific *perceptual vowel space*.

Special attention will be given to the *perceptual vowel space*. The relative psychological weight of specific acoustic parameters and the component dimensions of the perceptual vowel space are considered to vary across languages. Perceptual identification varies as a function of signal-inherent phonetic variation and the language-specific psychological weighting of specific acoustic dimensions.

Empirical evidence from a vowel identification experiment with L2 German learners from ten different native languages will be presented to illustrate how the same acoustic input may be identified as a token of one or the other phonemic vowel category.

To compare vowel systems and relationships between vowel categories, a further differentiation of *perceptual similarity* (psychological) vs. *articulatory or acoustic similarity* (physical) is necessary. Increasing distance in phonetic substance (acoustic or articulatory form) is considered to increase distinctiveness and to decrease perceptual similarity in order to guarantee phonemic contrast (function).

To visualize the perceptual vowel space, "*perceptual vowel maps*" will be derived by means of Multidimensional Scaling (Shepard 1972, 1980) from the experimental identification data. Moreover, alternative visualizations of the articulatory space and the acoustic vowel space will be discussed that provide a more accurate representation of articulatory constellations and their acoustic consequences than the traditional IPA-chart.

References

Hillenbrand, J. 2013. Static and dynamic approaches to vowel perception. In: Morrison, G. & Assmann, P. (eds.). *Vowel inherent Spectral Change. Modern Acoustics and Signal Processing*. Berlin, Heidelberg: Springer, 9-30.

Ladefoged, P. 1993. *A Course in Phonetics*, 3rd edition. Fort Worth: Harcourt Brace Jovanovich.

Shepard, R. 1972. Psychological representation of speech sounds. In: David, E. & Denes, P. (ed.). *Human Communication: A Unified View*. New York: McGraw-Hill, 67-113.

Shepard, R. 1980. Multidimensional Scaling, Tree-Fitting, and Clustering. *Science*, 210, 390-398.

Wood, S. 1975. The weakness of the tongue-arching model of vowel articulation. *Working Papers Dept. of Linguistics, University of Lund*, 11, 55-108.

Wood, S. 1979. A radiographic analysis of constriction location for vowels. *JPhon*, 7, 25-43.

Key words: phonetics – vowel space – similarity – multilingual mind – perception – form-function