

Contrast and the Feature Representation of Segments
(Oral or poster)
San Duanmu
University of Michigan
duanmu@umich.edu

How many vowels (or consonants) are there in the world's languages? Can we identify segments in different languages as being the same, even if they are not exactly so phonetically? Should we determine the features of segments (a) in terms of contrast or (b) in terms of natural classes? Would methods (a) and (b) yield different results? This talk addresses these questions through the analysis of two large segment corpora, UPSID (Maddieson & Precoda 1990) and P-base (Mielke 2004-2007).

It is well known that segments represented by the same IPA symbols can differ from language to language. For example, [i, y, e, ø] in Norwegian are quite different from those in German (Disner 1983). The German vowels generally have a higher tongue position than their Norwegian counter parts. In order to capture such fine differences between languages, some scholars (e.g. Ladefoged 1972, 1992; Flemming 2001, 2004; Port & Leary 2005) argue that it is insufficient to assume a set of binary, abstract, and universal features (e.g. Jakobson *et al.* 1952 and Chomsky & Halle 1968), as is commonly done in generative phonology. Instead, features should be gradient or scalar, fine grained, and language specific. On this view, we can compare phonetic similarities between segments in different languages, but there is little point identifying the segments as being the same.

Within a language, features are based on contrast: differences that can result in a lexical contrast must be represented with different features or feature values, and those that do not are not. I propose that the same method be applied to cross-linguistic comparisons: If the difference between X (a segment from language A) and Y (a segment from language B) can result in a lexical contrast in any language (say in language C), then X and Y should be represented with different feature values; if the difference between X and Y never results in a lexical contrast in any language, then there is no need to distinguish X and Y in terms of features.

According to the proposal, features are determined not only by looking at individual languages but also by looking at all languages. For example, if two segments X and Y contrast in the backness of the tongue, we still need to find out how many degrees of contrast in backness are possible in the world's languages. If the maximal number of degrees is two (Chomsky & Halle 1968), then the difference between X and Y is simply front-back. However, if the maximal number of degrees is three (Maddieson & Precoda 1990; Parker 2000), then the difference between X and Y could be front-back, front-central, or central-back.

This talk reports an analysis of vowels in UPSID and P-base. It is found that no vowel feature has more than a two-way contrast. This means that the maximal number of simple vowels is 24 (Chomsky & Halle 1968; Kiparsky 1974), rather than 42 (Maddieson & Precoda 1990). The result makes it possible to compare segments and features cross-linguistically, and to determine the frequencies or markedness of different segments and features. Frequencies based on contrast will be compared with frequencies based on natural classes (Mielke 2008) and possible sources of the differences will be discussed.