How universal is the boundary between vowels and consonants?

Classification into vowels and consonants is one of the most basic distinctions in phonetics. The IPA Handbook (1999, p. 6) describes it as “fundamental to the framework underpinning the IPA”. More recently segmentation into vocalic and consonantal intervals has been taken as a basis for studying rhythmic differences between languages (cf. Ramus 1999, Grabe 2002, among others). This approach rests on the assumption that speech can been presented as a sequence of vocalic and intervocalic intervals defined on a purely acoustic basis without reference to linguistic knowledge.

In this paper, we tested this assumption by measuring the extent to which familiarity with the language influences segmentation into vocalic and consonantal intervals. We have obtained manual segmentations of several texts in Modern Greek and Mandarin (ca. 1000 syllables) taken from the Oxford Aesop Corpus. For each paragraph we obtained several segmentations from fluent speakers of the language and several segmentations from phoneticians unfamiliar with the language. The latter were asked to segment the signal into vowels, consonants and sonorants, but had no access to the text or its phonological transcription.

We then used several methods to evaluate the inter-rater agreement amongst phoneticians who know the language, phoneticians who do not know the language and between phoneticians who know the language and those who do not.

The results showed that phoneticians familiar with the language had the highest agreement. The agreement between phoneticians not familiar with the language was also high. Yet the agreement between the two groups was lower than the agreement within each groups. This shows that while there is certainly substantial agreement about what is a consonant and what is a vowel, in many cases the decision depends upon familiarity with the language and the knowledge of lexical form. Listeners who know the language take advantage of a top-down process using linguistic/lexical knowledge of the language to classify vowel and consonants, while listeners who are unfamiliar with the language rely on a bottom-up approach using acoustic information alone. There are discrepancies between the linguistic information and the acoustic information, therefore, the labeler agreement is high within group but low across groups.

As one may expect, familiarity with the language had effect on the classification of approximants/short high vowels (cf. Ladefoged and Maddieson, 1996, p. 322) and syllabic consonants. The difficulties related to classification of such sounds have long been acknowledged and partially resolved by distinguishing between phonological vowels and consonants and phonetic vocoids and contoids (cf. Pike 1943). However, our results revealed further disagreements in classification that raise the question whether it is possible to separate phonological and phonetic classification. We conclude by discussing implications for cross-linguistic studies and phonetics in general.