A preliminary assessment of approaches to measuring phonetic similarity

How can we map our intuitions about speech sound similarities and dialect similarities onto numbers? What happens when individual vowel/consonant measurements are brought together to give an overall score of similarity between two dialects? Do we need a maximally phonetically detailed account or will a more sparse representation suffice? Beginning to tackle these three key questions not only has the potential to increase our understanding of the phonetics/phonology interface and dialectology, but also has relevance for historical linguistics and clinical linguistics (e.g. quantifying degrees of divergence from ancestral forms and target pronunciations respectively). The third question above (on level of detail) has sparked some sharp debate, with some preferring maximally detailed, linguistically informed analyses (e.g. Heggarty et al 2005), and others arguing that a sparser representation is not only sufficient but also necessary for statistical purposes (Kessler & Lehtonen 2006). Yet it is unclear which side has the upper hand, especially as Heeringa's (2004) comparison of measures incorporating vastly different levels of detail (e.g. minimally detailed whole phone comparison versus more detailed feature systems) led to remarkably similar results emerging them. I hypothesise that this levelling of differences may be due to the effect on phonetic detail of aggregating individual pairwise segment/word comparison scores into overall dialect scores, an issue on which there has been almost no discussion to date in the literature.

In this paper, after reviewing prominent approaches to phonetic quantification, I will contrast the whole phone comparison with two different articulatory feature systems (similar to Heeringa 2004), using a small subset of the English dialect data of McMahon et al (2005-07). The focus will be on the transition from segmental to dialect comparison, through which I will examine problems of standard methods of aggregation (e.g. mean scores) in a dialect context and propose alternatives. I will then outline my proposed next stages of a more full-scale comparison of the methods of Heeringa and Heggarty, which despite their representing strongly opposing schools (e.g. see Heggarty et al 2005), have never actually been directly compared against each other. These two elements represent the first strand of my empirical work. The second strand, of which I will also give brief details, is my attempt to construct an artificial data set using patterns in phonetic typology. The aim is to use this as a broader and more neutral platform for comparing different measures of phonetic similarity. I will conclude by mentioning the exciting possibility (for my future work) of incorporating prosodic factors into measurements, instead of being restricted to comparisons within isolated words, as is the case at present.

References

Heeringa, Wilbert. 2004. *Measuring dialect pronunciation difference using Levenshtein distance*. Groningen: University of Groningen Doctoral Dissertation.

Heggarty, Paul, April McMahon & Robert McMahon. 2005. From phonetic similarity to dialect classification: A principled approach. In Nicole Delbecque, Johan van der Auwera & Dirk Geeraerts (eds). *Persepctives on variation: Sociolinguistic, historical, comparative*, 43-91. Berlin: Mouton de Gruyter.

Kessler, Brett & Annukka Lehtonen. 2006. Multilateral comparison and significance testing of the Indo-Uralic question. Peter Forster & Colin Renfrew (eds.). 2006. *Phylogenetic methods and the prehistory of languages*, 333-42. Cambridge: McDonald Institute for Archaeological Research.

McMahon, April, Warren Maguire & Paul Heggarty. 2005-07. Sound comparions: Dialect and language comparison and classification by phonetic similarity. http://www.soundcomparisons.com/ (Mar 2008)