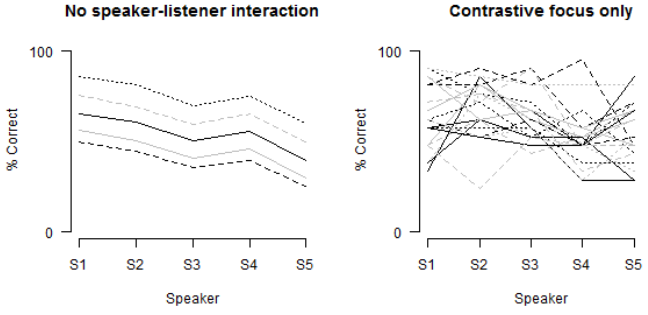
Listener-specific perception of speaker-specific productions in intonation

Francesco Cangemi, Martine Grice
IfL Phonetik, University of Cologne
{fcangemi,martine.grice}@uni-koeln.de

§1. Individual-specific behaviours for phonological contrasts at the segmental level have long been documented in production and perception. The interaction between the two – that is, the hypothesis that an individual's identification or discrimination performances are related to her articulation strategies or acoustic output – has been explored with the aim of supporting frameworks which posit a link between production and perception, yielding contrasting results (Bell-Berti et al. 1979, contra Paliwal et al. 1983). More recently, the issue has been explored in relation to speaker intelligibility (Newman et al. 2001, Hazan et al. 2013), and as a source of support for models of speech production in which articulatory movements are planned primarily in auditory space (Perkell et al., 2004).
§2. Beyond the segmental level, considerably less research has been devoted to speaker- and listener-specific behaviours in intonation (but see for production Niebuhr et al., 2011), and none to the interaction between the two. In this contribution, we document such an interaction in the encoding and decoding of pitch accent categories in German. In the production study, articulatory and acoustic recordings from 5 speakers uttering 7 repetitions of 4 target sentences in 4 dialogue contexts eliciting different focus interpretations (out of focus, broad, narrow and corrective focus) were collected (560 items). Speakers were found to vary with respect to the use of different acoustic cues (f0 peak alignment, f0 range, duration of target word) across the four focus conditions. In the perception study, 20 subjects provided forced-choice identifications through context-matching of 420 utterances from the production study (8400 items), evenly distributed across speakers, sentences and focus conditions.
§3. As could be expected, results pooled across focus conditions and examined through likelihood ratio tests on mixed effects models show that (i) some speakers are perceived better than others and (ii) some listeners perceive better than others. The main finding is that (iii) some listeners perform better on particular speakers, whereas other listeners perform badly on these speakers, but better on other speakers. Thus, results reveal a clear interaction between speaker-specific and listener-specific behaviours (Fig. 1 for one of the four focus conditions). Even more interestingly, (iv) the interaction between speaker-specific and listener-specific effects is modulated by focus condition, e.g. corrective focus items show considerably fewer listener-speaker interaction effects than broad focus items.
§4. These findings not only extend the available knowledge on production-perception links and speaker intelligibility beyond the segmental level (see §1), they are also particularly relevant to the issue of categories in intonation. Building on dynamic modelling (e.g. Gafos and Benus, 2006) of the relations between phonetics and phonology (e.g. Pierrehumbert et al., 2000), recent approaches to categoriality in intonation have stepped away from a monothetic perspective and suggested a distributional view (Nava, 2010). In this perspective, phonetic realizations of phonological categories are distributed along a continuum. Our finding that the strength of speaker-listener interactions depends on the intended focus structure (§3.iv) might help explain the differences in shape of distributions for the various categories, pointing to the existence of more compact vs. more diffuse categories in intonation.

Figure 1: Hypothesized results with no interaction (left) and actual results (right); each line represents one listener



References

Bell-Berti, F. et al. 1979. Some relationships between speech production and perception. *Phonetica* 36, 373-383. / Gafos, A., Benus, S. 2006. Dynamics of phonological cognition. *Cognitive Science* 30(5), 905-943. / Hazan, V. et al. 2013. The impact of variation in phoneme category structure on consonant intelligibility. Proc. ICA Montreal 19, 1-6. / Nava, E. 2010. *Connecting phrasal and rhythmic events: Evidence from second language speech*. PhD thesis, USC. / Newman, R.S. et al. 2001. Perceptual consequences of within-talker variability in fricative production. *JASA* 109, 1181-1196. / Niebuhr, O. et al. 2011. Are there "shapers" and "aligners"? Individual differences in signaling pitch accent category. *Proc. 17th ICPhS*, 120-123. / Paliwal, K.K. et al. 1983. Correlation between production and perception of English vowels. *J. Phon* 11, 77-83. / Perkell, J.S. et al. 2004. The distinctness of speakers' productions of vowel constrasts is related to their discrimination of the contrasts. *JASA* 116, 2338-2344 / Pierrehumbert, J. et al. 2000. Conceptual foundations of phonology as a laboratory science. In: Burton-Roberts, N. et al. (eds), *Phonological knowledge*. OUP, 273-303.