The perception of tonal targets in Neapolitan Italian: Tonal alignment and scaling.

Caterina Petrone

ABSTRACT.

Acoustic studies on tonal alignment claim the existence of striking regularities in the way in which fundamental frequency (henceforth F_0) peaks and/or valleys are aligned with the segmental string (the so-called "anchoring hypothesis") (Arvaniti *et al.*, 1998, Ladd *et al.,* 1999). However, tonal alignment can be affected by segmental structure (van Santen and Hirschberg, 1994; Rietveld and Gussenhoven, 1995). Even though research in various languages reports that the beginning (L) of a LH prenuclear rise is essentially timed at the onset of the stressed syllable, such constant alignment is not found for H(igh) targets at syllable offset. That is, the end of the rise has not a defined segmental landmark, but it is synchronized in different ways with the right edge of the stressed syllable or even with the following unstressed syllable. Recently, other studies on intonation also suggest that this alignment of tones to syllables is constrained by articulatory limitations such as speed of pitch change and phasing between coordinative structures (Xu, 2002), suggesting that these anchors can be universal. However there are few perceptual studies that test the significance of tonal alignment.

In the Neapolitan variety of Italian yes/no question LH nuclear rises are systematically later than (narrow focus) statement LH nuclear rises (D'Imperio 2000, 2001, 2002). Moreover, this regularity in production appears to be employed by listeners to distinguish such pragmatic contrast in perception (D'Imperio and House, 1997; D'Imperio, 2000). Moreover, the precise alignment of the H peak of both question and statement LH pitch accents varies as a result of syllable structure. For instance, the H peak of the L*+H of questions is timed to occur with the coda of the stressed syllable when the syllable is closed. We therefore hypothesized that Neapolitan Italian listeners capitalize on the alignment regularity for the perception of lexical contrast. Interestingly, we found that (D'Imperio, Petrone and Nguyen, to appear) tonal alignment produced a category boundary shift in CVC base stimulus series, suggesting that fine detail of tonal alignment is employed not only to signal pragmatic contrast but it might also be stored as part of the phonological specification of lexical items. That is, listeners employ the information conveyed by intonation in identification of minimal pairs.

Furthermore, even in acoustic analyses the precise tonal target is sometimes difficult to detect. For example, the tonal target can be masked by segmental perturbations: In this case, how do listeners interpolate the F_0 contour, computing both macro and micro-intonation in order to recognize the speech signal?

The tonal phenomenon of downstep (within the domain of downtrend phenomena) adds complexity to the definition of tonal targets. Downstep was originally used to define a typical phenomenon of scaling of F_0 in African tone languages, that is a stepwise lowering of immediately successive melodic peaks in an utterance (for a review, cf. Ladd, 1996) often triggered by a preceding sequence of tones. In Neapolitan Italian, the nuclear pitch accent of broad focus statements is H+L*. This accent seems to be less prominent than the prenuclear H* found in the same variety. Moreover, the pitch target is often realized as a plateau, making it even difficult to determine the timing of the peak. However, in perceptual experiments listeners have no difficulty in perceiving the prominence of this accent. It has been suggested that this accent could be downstepped, but the status of this accent is still unclear since it does not seems to be triggered by any determined tonal environment. Moreover, in Neapolitan Italian, a process of downstep seems to have been found also for accents in postfocal position, both in early focus question and in narrow focus statements. Even in this case the accent is still perceptually detectable by listeners (D'Imperio, 2002), though less prominent. Do downstepped tonal targets behave differently than non-downstepped accents? Do listeners have more difficulty in determining tonal alignment in the

perception of plateaus, given that the peak is not immediately discernable in the acoustic signal? All these questions are still open.

Future research in intonation will shed light on the contribution of information conveyed by F_0 and on interaction between production and perception mechanisms in speech. Specifically, we can determine to what extent do listeners have an implicit knowledge of the detailed consequences of phonetic process. Answers to these issues have obvious consequences not only for models of intonation but also for models of phonetic implementation and speech perception.