Focus, accentuation and phonetic variability in Greek
Argyro Katsika¹ and Amalia Arvaniti²
¹Haskins Laboratories ²University of Kent
¹argyro.katsika@haskins.yale.edu

Intonation has been investigated from two main perspectives, that of syntax and semantics, and that of phonetics and phonology. Research from the former perspective associates intonation (accentuation in particular) with information structure. This has often lead to the position that pitch accent identity, such as the contrast between H* and L+H* in English, is of little interest, on the assumption that what matters is distinctions in terms of information structure and focus type (such as contrastive, discourse-new and discourse-given; cf. Katz & Selkirk 2011). On the other hand, in the autosegmental-metrical framework of intonational phonology (AM), it is accepted that pitch accents cover a range of pragmatic nuances, with differences being understood to relate not only to information structure but to additional pragmatic concepts as well (e.g. Pierrhumbert & Hirschberg 1990; Roseano et al. 2016). At the same time, however, the importance attributed to differences in the scaling and alignment of pitch targets in establishing accentual contrasts leads to limited tolerance for variability in accent realization.

These two perspectives were examined with respect to the three accents used in focal position in Greek declaratives, H*, L+H* and H*+L (Arvaniti & Baltazani 2005). H* is used for discourse-new information. L+H* is used for contrastive focus. H*+L indicates that although the information is discourse-new, the addressee believes it should have been part of the common ground. Speakers (N = 13) read mini-dialogues designed to elicit polysyllabic test words bearing one of the three accents. In 18 of the dialogues the accents were examined as a function of phrase length (one-word vs. multi-word) and stress location; the test words were [laðoˈlemono] “oil-and-lemon-sauce”, [lemo naˈna] “lemonade”, and [yala na] “blue”. In another 9 dialogues [ˈmalama] “gold” was used to elicit the pitch accents in three tonal crowding conditions, with 0, 2, or 4 syllables between the focal accent and the one immediately preceding it. In all dialogues, test words were phrase final.

For each test word, the interval between the onset of the antepenult and the offset of the ultima (i.e. the offset of the word) was marked, and the F0 signal of this interval was extracted at 10 ms steps using Praat (segment identity ensured smooth F0 contours). The Lucero et al. (1997) nonlinear time warping technique was used to compute the normalized alignment of the F0 signals that corresponded to each condition, and the resulting averaged signals were compared across conditions.

Preliminary analysis based on two speakers shows systematic differences between the three accents in the alignment and scaling of the pitch peak (see Fig. 1). Further, the data show consistent differences in the realization of the tonal targets that depend on the greater (rather than just the immediate) context: e.g. in [laðoˈlemono], the reflex of the L tone in L+H* is clearly located at the onset of the stressed syllable in the one-word rendition (Fig. 1e) but not when other material precedes, even when there is no tonal crowding involved (Fig. 1b); such differences between the multi-word and one-word contexts apply to all test words. Further, the data show systematic variation that is not limited to tonal target scaling and alignment, but evidence non-localized effects on F0 as well as effects on other phonetic categories. Thus, differences are found first in the slope of the F0 fall after the peak, with L+H* and H*+L showing steeper after-peak slopes than H* (Fig. 1, all panels), and second in duration, with words accented with L+H* showing longer duration overall compared to their counterparts with H* or H*+L accent (results not shown). Similarly, while H* is realized as an F0 drop largely localized on the stressed syllable in [ˈmalama] and [laðoˈlemono], an F0 drop of steady slope extends throughout the word in [yala na] in which the last syllable must accommodate not only H* but also the L-L%d edge tones following it (compare Fig. 1, panels (a), (e) and (g)); this effect extends much further to the right than tonal crowding would predict (Arvaniti & Ladd 2009). Non-local effects consistent with the idea of an extended window for accent realization are also evident in overall contour shape, especially for H*+L (see Fig. 1, panels (c), (f), (d) and (g)). Finally the data reveal extensive variability in scaling across conditions (though relationships between accents remain by-and-large consistent), as well as substantial interspeaker variability (though differences between accents are comparable across speakers).

The systematic production of different F0 shapes in response to prompts that created different discourse conditions compatible with the pragmatic meanings of the accents discussed above supports the analysis of Greek as having three focal accents for declaratives, H*, L+H* and H*+L. As these three
accents do not correspond neatly to widely accepted focus types, this result supports the AM view that languages may encode a variety of pragmatic nuances using pitch accents, rather than being necessarily limited – as syntactic and semantic views of accentuation suggest – to distinctions between new, given and contrastive focus.

The results, however, also confirm that accents can be eminently variable across contexts and speakers. What is of particular interest here is the fact that using non-linear time-warping brings to the fore the non-localized nature of this variability, which extends to preceding and, where appropriate, following syllables (cf. the slope of the drop after L+H*). These non-localized effects, which will be quantified, are crucial for our understanding of the phonetics of pitch accents and the modelling of intonation more generally, and can have repercussions for the phonological representation of accents. From a phonological perspective, results derived from the kind of measurements employed here could help better determine the phonological identity of accents by separating essential from non-essential elements of an accent’s tonal composition (e.g. by examining resistance to variability); in turn, the significance of these elements for pragmatic processing can be further tested with perception experiments. From a phonetic perspective they show that simply measuring the scaling and alignment of turning points and pooling data across study participants is not sufficient to adequately describe intonational phonetics (cf. Barnes et al. 2012), and advocate for more research on both phonetic detail and the articulatory mechanisms that could be driving these global changes. They also argue in favour of using more varied samples of data rather than relying on a limited set of test words which may or may not be representative of the variability inherent in the realization of accents. Overall the results point towards a view of accents as distributions of values (in line with all phonetic categories) rather than as invariable prototypes or sets of discrete “allophones” as is often the practice in AM.

References


