Gesture apex coordination with prosodic structure and tonal events in Maltese English

Martine Grice[1], Alexandra Vella[2], Maria Lialiou[1], Florence Baills[3], Aviad Albert[1], Petra B. Schumacher[1], Nadia Pelageina[1] & Solveigh Janzen[1]
[1]*University of Cologne,* [2]*University of Malta,* [3]*Universitat de Lleida*
martine.grice@uni-koeln.de

This study investigates the co-occurrence of manual gestures with prosodic structure and associated tonal events in Maltese English (MaltE). In this variety of English, H tones can be on a metrically stressed syllable, in which case they constitute pitch accents, but, similarly to Turkish [1], tones can also occur on an initial unstressed syllable of a prosodic-word-sized constituent, which can be the initial syllable of a content word or a (monosyllabic) function word preceding the content word. When there is an initial H tone (an 'early peak' [2]), tone and lexical stress compete as cues to prominence, contributing to a reduced sensitivity to stress in sequence recall tasks ("stress deafness") in bilingual speakers of Maltese and MaltE [3].

It has been proposed that the strokes and apices of referential and non-referential manual gestures co-occur with pitch accents or, more broadly, the stressed syllables associated with them [4,5,6,7,8]. To shed more light on the prominence cueing potential and prosodic status of early peaks in MaltE, we ask whether these early peaks can function as gestural anchors in a similar way to pitch accents. To address this question, we analysed a 14-minute long TEDx talk in MaltE [9]. We annotated H* pitch accents and early H tones [10] and, in a separate step, gestural strokes and apices [11]. First, to assess the alignment between gestures and tones, we calculated the distance between the annotated gesture apices and the F0 peaks corresponding to H* and early H tones within the same content word (as well as H on preceding function words; H and H* were never on the same word). In Fig. 1, the distribution that links the gesture apex to a H* pitch accent peak is narrow and centred around 0 (blue), indicating a robust link between gestural apices and H*. Since our prosodic annotation was restricted to H and H*, this robust link covers less than half of the gestural apices in our data (202 out of 445). The distributions that link the gesture apex to an early H tone within the content word (pink) are broader and reflect even fewer cases. Moreover, when the H tone occurs on a preceding function word (green), the timing appears to be normally distributed around 250 ms after the H peak rather than being skewed leftwards to reflect any link between the gesture apex and the early H peak.

The picture changes when we switch our timing perspective from the earlyH/H* peaks to the lexically stressed syllable, with substantially more cases appearing to be linked in this way: regardless of the type of tonal event linked with gesture-syllable pairs, they all seem to be normally distributed within the stressed syllable (376 out of 445 apices are distributed around the centre of the stressed syllable). Fig. 2 shows the 272 cases that we could link to an early H or H* peak. Crucially, the subset of cases related to the early H on a preceding function word (green) have a similar distribution to those related to H* (blue). Early H peaks, however, do seem to slightly attract the gesture apex when occurring on the initial weak syllable of a content word: these results show a leftward trend relative to the stressed syllable (pink), indicating a possible additional coupling with the early H in this case [8].

We show that the co-occurrence of gestural and prosodic units contributes to our analysis of the intonational phonology of MaltE: the fact that gesture apices align differently with H* and early H provides evidence for the two tones being phonologically distinct categories. Moreover, MaltE is different from Turkish [1], in which the lack of a discernible H peak on the content word leads to the gesture apex being aligned with the beginning of the prosodic word. Our results indicate that gesture apices may not be directly linked to H tones in terms of turning points in the F0 curve, but to potentially tone bearing syllables, i.e. to metrically strong syllables. We are currently planning follow-up studies that will explore gestural apex co-occurrence with syllabic landmarks in the acoustic signal using the ProPer toolbox [12].

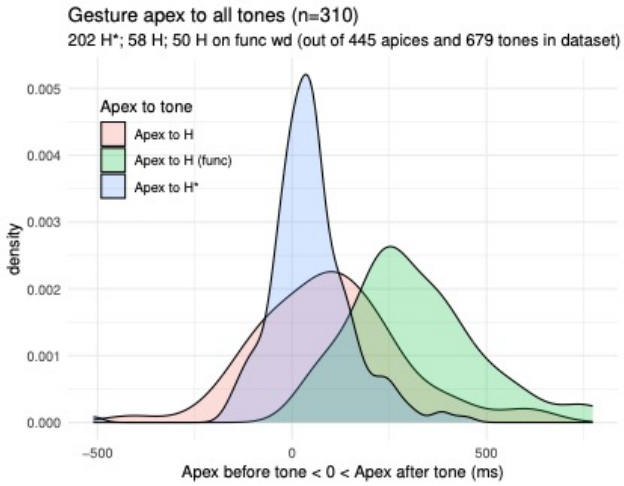

Figure 1: *Alignment of gesture apex with F0 peaks corresponding to H*, early H (on content word) and early H (on preceding function word).*

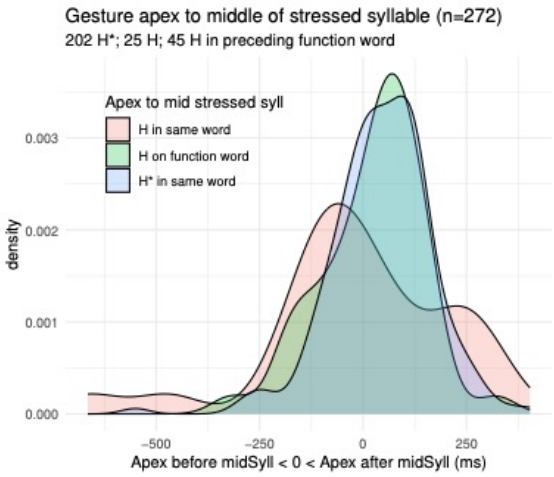

Figure 2: *Alignment of gesture apex with centre of stressed syllable when there is a H* on this syllable, and when there is an early peak on the same (content) word or on a preceding function word.*

References

[1] O. Turk and S. Calhoun, "Multimodal cues to intonational categories: Gesture apex coordination with tonal events", *Laboratory Phonology* 14(1), 2023.

[2] A. Vella, "Alignment of the 'early' HL sequence in Maltese falling tune wh-questions", Proc. 15th International Congress of Phonetic Sciences pp. 2062-2065, 2011.

[3] M. Lialiou, A. Bruggeman, A. Vella, S. Grech, P. B. Schumacher and M. Grice, "Word-level prominence and 'stress deafness' in Maltese-English bilinguals", Proc. 20th International Congress of Phonetic Sciences, pp. 132-136, 2023.

[4] S. Shattuck-Hufnagel and A. Ren, "The prosodic characteristics of non-referential co-speech gestures in a sample of academic-lecture-style speech", *Frontiers in Psychology* 9, 2018.

[5] P. Rohrer, E. Delais-Roussarie and P. Prieto, "Visualizing prosodic structure: Manual gestures as highlighters of prosodic heads and edges in English academic discourses", *Lingua* 293, 2023.

[6] A. Gregori and F. Kügler, "The distribution of co-speech gestures, information structure and prosody: A corpus study on prominence peak alignment" Talk at Phonetik und Phonologie im deutschsprachigen Raum. https://doi.org/10.11576/pundp2022-1029, 2022.

[7] S. Repp, L. Muhtz and J. Heim, "Alignment of beat gestures and prosodic prominence in German", Proc. Interspeech, 107–111, 2023.

[8] H. R. Bosker, M. Hoetjes, W. Pouw, W. and L. van Maastricht, "Gesture-speech coupling in L2 lexical stress production: A pre-registration of a speech acoustic and gesture kinematic study". 2021, https://doi.org/10.31219/osf.io/w2ezs

[9] TEDx University of Malta: "Till death do us part", https://www.youtube.com/watch?v=SVHda2x4B6Y

[10] A. Amalia, and J. Fletcher, "The Autosegmental-Metrical Theory of Intonational Phonology", in C. Gussenhhoven, and A. Chen (eds), *The Oxford Handbook of Language Prosody*, Oxford Handbooks, 2020.

[11] P. Rohrer, U. Tütüncübasi, I. Vilà-Giménez, J. Florit-Pons, G. Gurrado, N. Esteve-Gibert, A. Ren-Mitchell, S. Shattuck-Hufnagel and P. Prieto, "The MultiModal MultiDimensional (M3D) labeling system". doi: 10.17605/osf.io/ankdx/

[12] A. Albert, F. Cangemi, T. M. Ellison, and M. Grice. ProPer: PROsodic analysis with PERiodic energy [Computer software], 2023. https://osf.io/28ea5/