Somewhere over the spectrum: between robotic and singsong intonation

Francesco Cangemi, Simon Wehrle, Martina Krüger, Martine Grice

IfL Phonetics, University of Cologne

Intonation styles, resulting in the voice sounding robotic or singsong, have been characterised as a property of individuals, of specific groups of individuals, of language varieties, speech register and non-native speech. At the level of the individual, intonation styles are important for forensic applications and, together with voice quality, for personality assessment (Mohammadi et al. 2012). The intonation style of specific speaker-groups, such as people with Autism Spectrum Disorders (ASD), have been characterised as both robotic (Kanner, 1943; Green and Tobin 2009) and sing-song (Simmons & Baltaxe, 1975), or both (Baltaxe, 1984). Studies in perceptual dialectology often find sing-songiness to be highly relevant in subjects' judgements, e.g. Kuiper (1999) shows that Parisians perceive the southern Provençal variety to be "sing-song" and "singing" (see also Nolan 2006). The use of sing-song intonation is also related to speech register, as this style is said to characterise not only infant directed speech (e.g. Holmes 2013) but also speech by adults talking to attractive conversation partners (Leongómez et al. 2014). Finally, both intonation styles play a role in the characterisation of second language pronunciation. Speakers of Dutch are said to have a narrower f0 range in their L1, compared to speakers of Swiss German (Celce-Murcia et al., 1996), a difference that surfaces in the L2 English intonation of Dutch-accented English, which sounds "somewhat flat", and in the productions of Swiss-German natives, who have "a somewhat sing-song quality to their English" (Celce-Murcia et al. 1996: 193).

Despite its potential for linguistics, the characterisation of intonation as robotic or sing-song remains ill-defined, reflecting the still rather rudimentary nature of analyses of pitch range, a crucial component in the characterisation of intonation styles (see Patterson 2000, Mennen et al. 2012, Urbani 2013). In these studies, emphasis is placed on *long-term distributional* (LTD) measures such as the range, mean, skewness and kurtosis of f0 contours, and on *linguistic* measures, where landmarks in the f0 contour are identified on the basis of their proposed linguistic significance. With this contribution, we aim to complement this picture by characterising intonation behaviour *dynamically*, focussing on f0 trajectories across time. We characterise utterances on the basis of:

(i) *Slope Changes*: the number of inflection points in the f0 contour

(ii) *Max Excursions*: the extent of the largest f0 movements

Specifically, f0 contours are extracted with *Praat* (Boersma & Weenink 2017) and hand-corrected with *mausmooth* (Cangemi 2015). The contour is then stylised with a resolution of 2 semitones. The number of turning points thus extracted is further normalised by the duration of the utterance, yielding the measure *Slope Changes*. This measure represents how often the f0 contour changes slope across the utterance, and is expected to be higher for sing-song contours. Next, the f0 distance between each pair of two successive turning points is calculated, and the two highest distances are averaged into the measure *Max Excursions*. This measure represents how large f0 movements can be, and is expected to be higher for sing-song contours. Since f0 contours can be both more or less wiggly (*Slope Changes*) and more or less spacious (*Max Excursions*), the two measures are largely independent, and are thus used to provide a dynamic account of sing-songiness in intonation.

Figure 1 provides an informal application of this method to the intonation of people with high-functioning ASD. Each point is an utterance spoken by either one of two adult German female speakers with ASD (black and grey circles) or one neurotypical control speaker (empty circles), collected in a Map Task context (Brown et al. 1983). ASD speakers are colour-coded; grey dots represent utterances from a speaker informally perceived as being more sing-song. As expected, they are characterised by both higher *Slope Changes* and higher *Max Excursions*. Interestingly, the figure points to the impossibility of characterising the speech of adults with ASD as a single group. This helps to explain the contradicting results found in the literature on the topic, and suggests that each individual with ASD might be placed along a spectrum that goes from robotic to sing-song, with the values of control speakers in between these two poles.

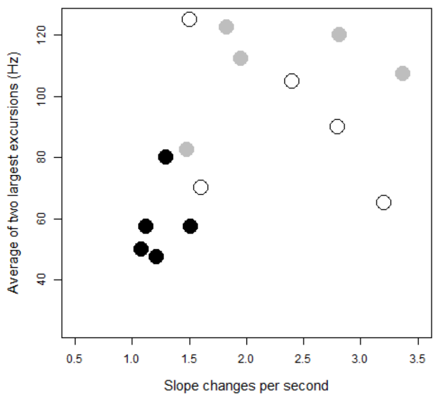

Figure 1

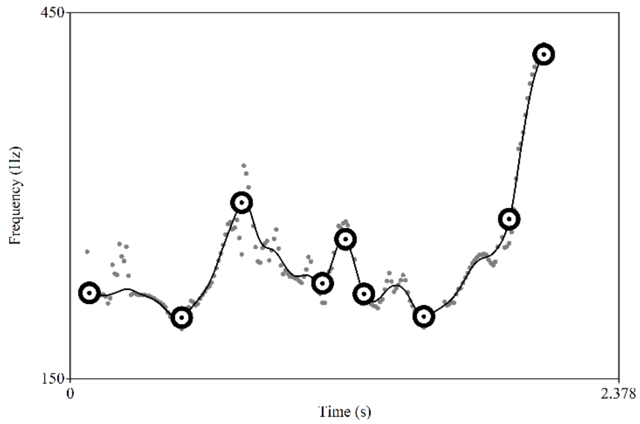

Figure 2

Figure 2 shows the f0 contour of a single test utterance, specifically the one represented as the rightmost grey dot in Figure 1 (around coordinates *x=3.3* and *y=105*). The extracted points are plotted with small grey speckles, the "mausmoothed" contour is plotted with a thin black line, and the stylised turning points are plotted as dots within circles. With 8 turning points (after the first) in a total time of ~2.3sec, the *Slope Changes* is ~3.3 changes per second (see *x* coordinate). With the largest f0 movements being the one between the last two points (~135Hz) and the one immediately preceding it (~80Hz), the average for *Max Excursions* is ~105Hz (see *y* coordinate).

Baltaxe (1984), in *Journal of Speech and Hearing Research*.
Brown, Anderson, Yule, Shillcock (1983), *Teaching Talk.* Cambridge University Press.
Cangemi (2015), *mausmooth*. Retrieved from author's webpage.
Celce-Murcia, Brinton, Goodwin (1996), *Teaching Pronunciation*. Cambridge University Pr.
Holmes (2013), *An Introduction to Sociolinguistics*. Routledge.
Kanner (1943), in *Nervous Child.*
Kuiper (1999), in Preston (ed.), *Handbook of Perceptual Dialectology.* Benjamins.
Leongómez, Binter, Kubiková, Stolarová et al. (2014), in *Evolution and Human Behavior*.
Mennen, Schaeffler, Docherty (2012), in *Journal of the Acoustical Society of America.*
Mohammadi, Origlia et al. (2012), in *Proceedings 20th ACM Int. Conference on Multimedia*.
Simmons, Baltaxe (1975), in *Journal of Autism and Childhood Schizophrenia*