DOI: 10.17469/O2104AISV000010

SIMON WEHRLE, FRANCESCO CANGEMI, MARTINA KRÜGER, MARTINE GRICE

Somewhere over the spectrum: Between robotic and singsongy intonation

The impressionistic characterisation of intonation as “robotic” or “singsongy” is frequent in many phonetics-related fields, such as forensic linguistics, clinical linguistics, perceptual dialectology and language acquisition. Despite its potential for linguistics, however, the characterisation of intonation as flat or sing-songy remains ill-defined. With this contribution, we propose a dynamic characterisation of intonation, focussing on trajectories of fundamental frequency (F0) across time. We apply this method to the issue of intonation in adults with autism spectrum disorders (ASD), which has variously been reported to be both more singsongy and more robotic than the intonation of neurotypically developed speakers. Our results point to the impossibility of characterising the speech of adults with ASD as a single group, thereby offering an explanation for previous contradictory results and highlighting the importance of individual variability.

Keywords: intonation, sing-song, robotic, pitch range, autism spectrum disorder.

1. *Introduction*

1.1 Overview

At first glance, judging a speaker’s intonation style seems to be a relatively straightforward task. Listeners intuitively form impressions based on intonation, amongst other things, in many different contexts and without conscious effort. Putting such impressions into words with any degree of accuracy and confidence is a much more difficult task, however, often resulting in the use of a very limited range of terms, such as “robotic” (i.e. monotonous) or “singsongy” (i.e. lively and repeatedly spanning a large range), at the two ends of the scale. An even greater challenge lies in the formation of scientifically testable operationalisations and the choice of appropriate measurements in an effort to uncover the underlying mechanisms and parameters of intonation styles.

In this contribution, we present a new method of measurement which is shown to be capable of reliably quantifying intonation styles. We exemplify this approach using data from the speech of subjects diagnosed with autism spectrum disorders (ASD). Speech in ASD has in the past been described using *both* extremes of characterisation mentioned above (robotic and singsongy), reflecting the problems inherent in relying on such vaguely defined and technically underspecified terms.

We suggest that a further reason for the contradictory claims on intonation in ASD lies in the infelicitous practice of relying on averaged values across groups of subjects without due consideration of variation at the level of the individual. The importance of considering scientific data at this level is not specific to this study (cf. Cangemi, Krüger & Grice, 2015). It is, however, made all the more critical when trying to understand the behaviour of a group of speakers as heterogeneous as that of individuals diagnosed with ASD. Considering speaker-specific data becomes nothing less than a necessity when we are additionally dealing with (very) small sample sizes of a population, as has been the case in the vast majority of studies dealing with speech in ASD. Our data seem to show that an inappropriate reliance on mean values across speakers has to be considered as an underlying reason for the conflicting findings describing the intonation style of speakers with ASD as either robotic or singsongy.

1.2 The linguistic interest of intonation styles

Intonation styles in general are of interest to linguists for a variety of reasons. First, they are a property of individual speakers. Besides the character attributions formed in everyday spoken interaction, this facet of individual specificity is of interest from both a more practical and a more theoretical standpoint. Practical applications include forensic phonetics and emotion profiling (Ladd, Silverman, Tolkmitt & Scherer, 1985; Mohammadi, Origlia, Filipponi & Vinciarelli, 2012). Regarding theory, the issue is pertinent both to the long-standing debate on idiolects (Paul, 1880) and to the more recent debate about the concept of individual grammar networks (Cangemi et al., 2015).

Intonation styles are relevant, too, for describing the behaviour of groups of individuals. Intonation has featured particularly prominently in research on the speech of one such group, people with autism spectrum disorders. Speakers with ASD are generally said to have "atypical" intonation. Quite what this means and how it can be measured is less clear and what emerges from the limited number of studies investigating this phenomenon is far from conclusive. It has been suggested since Simmons and Baltaxe (1975) that people with ASD often use a singsongy intonation with excessive pitch variation. However, flat and robotic-sounding behaviour has been documented since Kanner (1943; see also Green & Tobin, 2009). Both of these findings are consistent with Baltaxe (1984), who, intriguingly, showed that autistic children had either a very narrow F0 range or a very large F0 range.

Moving beyond groups of individuals, intonation styles are no less relevant for the description of language varieties. There is abundant evidence for the importance of intonation styles for impressionistic judgements of different dialects: Data from Kuiper (1999: 258) shows that Parisians consider the southern Provencal variety to be "singsongy" and "singing", while they consider the eastern Alsatian variety to be "choppy" and "jerky" (see also Nolan, 2006). It is hard to disentangle such attributions from the wide range of cultural factors and stereotypes that may play a role, but speech styles in themselves are sure to be one crucial factor underlying such descriptions. Such speech styles in turn are influenced by the phonological properties of the regional variety spoken. For instance, the varieties of French spoken in the

South are characterised by final schwas (Coquillon, Durand, 2010). This extension of segmental material available for the production of intonation contours provides an opportunity for more pitch movement (cf. Torreira, Grice, 2018), which, while it does not necessarily have to lead to a more lively intonation, certainly could be one factor underlying the impressions of sing-songiness cited above.

The use of singsongy vs. flat intonation also seems to be related to choices in register, as, for instance, sing-songiness is characteristic of infant-directed speech (IDS) (e.g. Holmes, 2013). A more exaggerated, "motherese" speech style has been shown to lead not only to better mother-infant bonding, but also higher intelligibility and, consequently, better later language development in infants (Liu, Kuhl & Tsao 2003; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008). Lively F0 use furthermore characterises speech by adults talking to attractive conversation partners (Leongómez, Binter, Kubicová, Stolarová, Klapilová, Havlíček & Roberts, 2014). Why a more singsongy intonation is used with interlocutors of a greater attractiveness is not entirely clear, but while probably not being orthogonal to experiences of, and positive associations with, IDS, it might also reflect evolutionally desirable traits such as liveliness and lack of threat. Decreased F0 variability has, conversely, been reported as characteristic of competitive contexts with high aggressiveness (Hodges-Simeon, Gaulin & Puts, 2010).

Finally, intonation styles are relevant for bilingual and second language speech. It has been suggested that different languages can have narrower or larger F0 ranges overall. For instance, Dutch and Japanese have been described as having a narrower F0 range than English; Swiss German and Norwegian have been described to have a wider F0 range than English (Celce-Murcia, Brinton & Goodwin, 1996; Graham, 2014). Celce-Murcia et al. (1996) project data describing F0 range in several native languages onto English as a second language. The described differences of F0 range in the L1 are said to be reflected in the L2, with e.g. Dutch-accented English being described as sounding "somehow flat" and Swiss German-accented English described as having "a somewhat sing-songy quality" (ibid: 193).

Despite the relevance of intonation styles to manifold aspects of language and their being a phenomenon of interest at various levels of linguistic inquiry, the associated methods of measurement have been far from uniform. Perhaps surprisingly, no dedicated attempt that we are aware of has been made to tackle the issue of how intonation styles can be quantified appropriately. The work presented here aims to remedy this situation.

1.3 Intonation styles and pitch range

The precise nature of intonation styles beyond subjective characterisation has remained ill-defined, but there is a long tradition of studies investigating the closely related concept of pitch range (see Lehiste, 1975; Ladd et al., 1985). As a complement to the long-established measurement of mean F0 in the description of prosody, various approaches have been made in order to capture what are essentially the levels and variations of a speaker's minimum and maximum pitch. The most recent and widespread characterisation of pitch range can be found in the work of Mennen, Schaeffler & Docherty (2012) and subsequent work by e.g. Urbani (2013) and Graham (2014). This approach is based on the assumption that pitch range is best described through a combination of linguistic and distributional parameters. In the following, we will further examine this method and point out how we think it may be complemented and refined in order to better capture different styles of intonation.

Ladd, Terken (1995) and Patterson (2000) first suggested using what they call "linguistic measures" in order to determine pitch range. This entails identifying "linguistically relevant landmarks" (Mennen et al., 2012) in the F0 contour and using them, rather than global minima or maxima, to characterise a speaker's F0 range. In practice, the F0 contour is reduced to a series of high or low turning points, which are then labelled and averaged (within equivalent labels). This approach has shown convincing results in its application to a number of languages, but parts of it still leave room for improvement. For instance, the basis for the chosen operationalization is not driven by theoretical deliberations, but rather by pragmatic reasons, as pointed out by the authors themselves:

> Our decision to assume a direct relationship between turning points and phonological tones was driven by practical reasons so as to ensure consistency in our labelling. However, tones and turning points may not necessarily map in a one-to-one fashion, so that some tones may not be realized as turning points and some turning points may not constitute an underlying phonological tone (ibid., footnote 3).

More importantly, the meaning of intonational labels in itself has come under increasing scrutiny and critical re-examination in recent years (see the contributions in D'Imperio, Grice & Cangemi, 2016). In the method for measuring pitch range outlined above, intonational labels are taken as the *starting point* for further analyses, providing a symbolic reduction of the phonetic signal. This is consistent with a widespread approach in intonation research (from Hirst, Di Cristo, 1998, to Hualde, Prieto, 2016). However, recent developments suggest that it could be more fruitful to take the opposite approach and use intonational labels only as the *outcome* of phonological analysis (Cangemi, Grice, 2016; Frota, 2016). In this perspective, the use of labels requires an evaluation of intonational meaning and of prosodic structure, rather than a discretisation of the phonetic signal.

Besides turning points based on symbolic labels, the second pillar of the approach by Mennen et al. takes the form of so-called "Long-Term Distributional" (LTD) measures. These measures essentially comprise the range, mean, skewness and kurtosis of the distribution of F0 values. Although useful for descriptions of pitch range, the reason why LTDs are nevertheless not ideal for exploring intonation styles can be illustrated with the following example. Consider the F0 contour in Fig. 1. Whilst this contour may not be something that will ever be found in human speech data, it is a useful idealisation of the shape an imagined F0 contour truly worth of the description "robotic" might take. To show why LTDs are problematic for the present purpose, compare the contour in Fig. 1, which is relatively monotonous (but mainly monotonic in the mathematical sense, i.e. entirely non-decreasing), with the one in Fig. 2, which represents the other end of the scale: an extreme version of a thoroughly lively, singsong intonation style. The problem is that these two very different contours yield exactly the same result in an analysis of LTD measures (see Fig. 3), thereby completely obscuring the essential difference between the two styles of intonation – at least in the hypothetical, stylised versions considered here. For this reason, LTDs along with linguistic measures based on phonological labels cannot, to our minds, be considered an entirely satisfactory measurement for the characterisation of intonation styles that is the aim of the present study.

Figure 1 - *Hypothetical F0 contour of a monotonic intonation style*

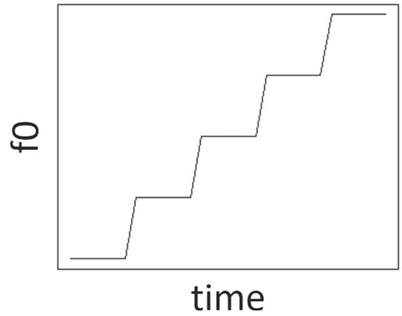


Figure 2 - *Hypothetical F0 contour of a singsongy intonation style*

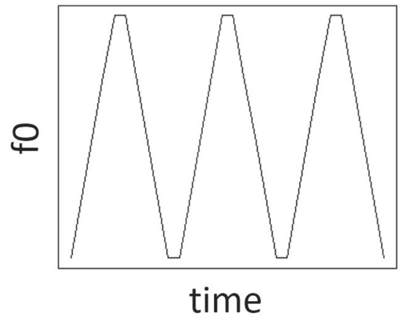
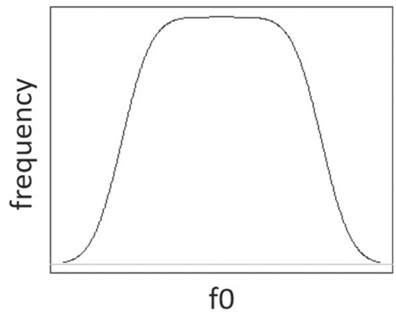


Figure 3 - *Frequency distribution (LTD) of both the monotonic F0 contour shown in Fig. 1 and the lively F0 contour shown in Fig. 2*



2. *Method: A dynamic characterisation of intonation styles*

The novel approach of characterising intonation styles presented in this paper aims to avoid the pitfalls inherent to an approach relying on linguistic and Long Term Distributional measures by instead focussing on the dynamics of F0 contours, represented in the time course of F0 trajectories. Two parameters that capture this aspect are presented in the following: Wiggliness and Spaciousness.

Wiggliness is operationalised as the amount of times an F0 contour "changes direction" over a given stretch of time, i.e. how many different rises and falls are contained within the portion of speech under investigation (based here on a stylisation of the F0 contour with a resolution of 2 semitones).

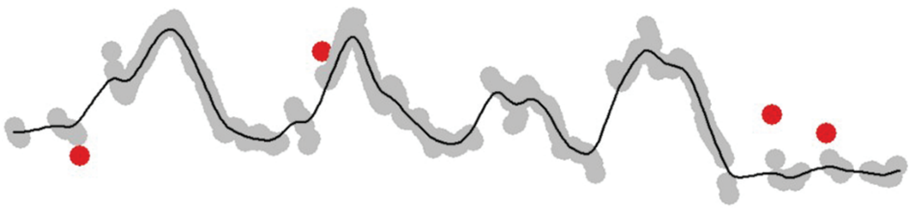
Spaciousness is operationalised as the extent of the slopes of these individual rises and falls, i.e. the maximum F0 excursions.

The more wiggly and spacious the contour, the more singsong we expect it to be and the less wiggly and spacious the contour, the more robotic we expect it to be. As F0 contours can be both more or less wiggly and more or less spacious, the two measures are at least partly independent and are thus chosen to provide a dynamic account of intonation styles.

In a demonstration of how to put this concept into practice, we first choose an excerpt of speech. The length of the excerpt is not fixed and can consist of e.g. one intonation phrase or one interpausal unit.

Next, the F0 contour contained within this excerpt is extracted in Praat (Boersma, Wennink, 2018) and semi-automatically corrected and smoothed using *mausmooth* (Cangemi, 2015). The *mausmooth* procedure is used to first identify any mistakes or artefacts in the Pitch object created by Praat. After correction or deletion of relevant cases, all remaining points are then transformed into a single smooth, continuous contour (see Fig. 4).

Figure 4 - *Correction and smoothing of an extracted F0 contour from an excerpt of speech using mausmooth. Grey dots represent the original pitch contour extracted in Praat, red dots represent points from this original extraction that have been manually corrected or deleted and the black line represents the final smoothed contour.*



In a next step, *Praat*'s Manipulation function is used to stylise the smoothed curve with a 2 semitone resolution (see Fig. 5). This smoothed and stylised curve is the input for further processing which will then yield the characterisation of intonation styles along the dimensions of Wiggliness and Spaciousness introduced above.

The threshold of 2 semitones for smoothing is used here as a first approximation of how the intonation contour might be perceived. By applying smoothing before stylisation, turning points are only located where an actual tonal movement is likely to be perceived. For this reason, we exclude from further analysis certain turning points which are visible in the F0 contour but which are not retained after both the smoothing and stylising procedures (such as the one indicated by the arrow in Fig. 5).

Figure 5 - *Stylisation of the smoothed F0 contour in Praat with a 2 semitone resolution. The arrow indicates an apparent turning point in the F0 contour which is not retained after smoothing and stylising*

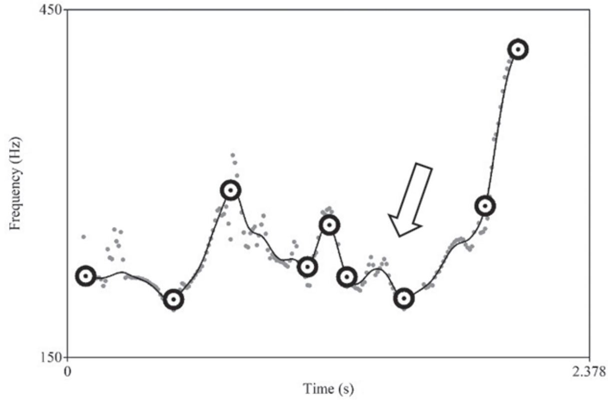
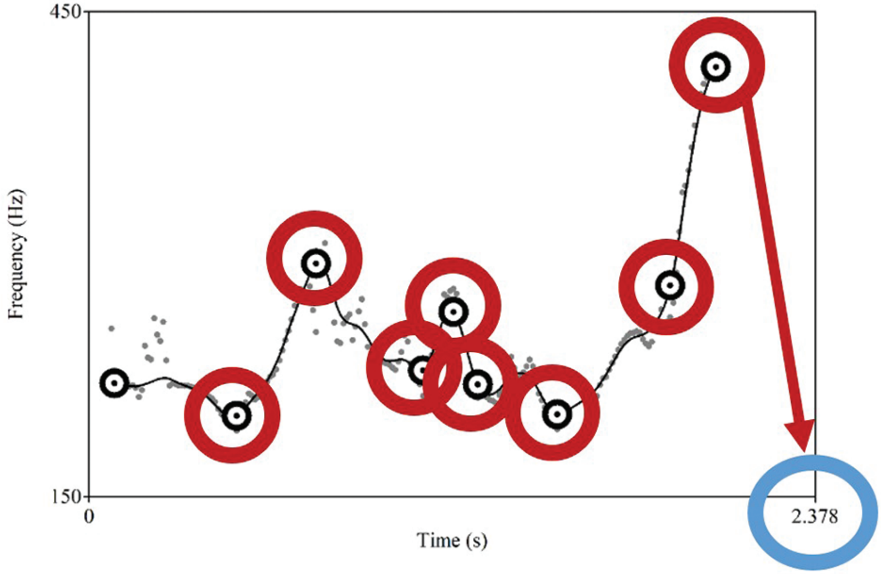


Figure 6 - *The measure of Wiggliness, or Slope Change, is obtained by counting the number of turning points in the stylised F0 contour and dividing it through the length of the excerpt in seconds. In this example, we have 8 turning points after the first one and a total duration of 2.378 seconds, yielding a Wiggliness measure of 3.364*

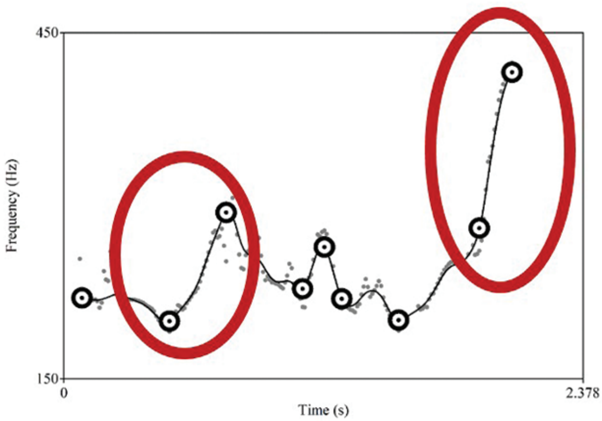


In order to obtain the measure of Wiggliness, or Slope Change, we simply count the number of turning points in the stylised curve and divide this number by the duration of the chosen excerpt in seconds (see Fig. 6).

In order to obtain the measure of Spaciousness, or Maximum Excursions, we simply identify the two largest F0 movements between two turning points and then calculate their average (see Fig. 7).

It is worth pointing out that neither the choice of a 2 semitone resolution for stylisation nor the choice of precisely the 2 largest F0 movements to obtain the averaged value for Spaciousness are extrinsically or theoretically motivated, but simply reflect a starting point for exploration that has proven successful for our data so far. The exact values of these parameters can be adapted and fine-tuned in future work depending on the speech material under investigation. Furthermore, results gained from perception studies designed to test for the perceptual relevance of the measures proposed here will either corroborate or refute their usefulness and guide subsequent refinement of these values.

Figure 7 - *The measure of Spaciousness, or Maximum Excursions, is obtained by identifying the two largest F0 movements between two turning points and then calculating their average. In this example, the two largest excursions have a value of approximately 80 Hz and 135 Hz, respectively. This yields a Spaciousness measure of approximately 105 Hz*



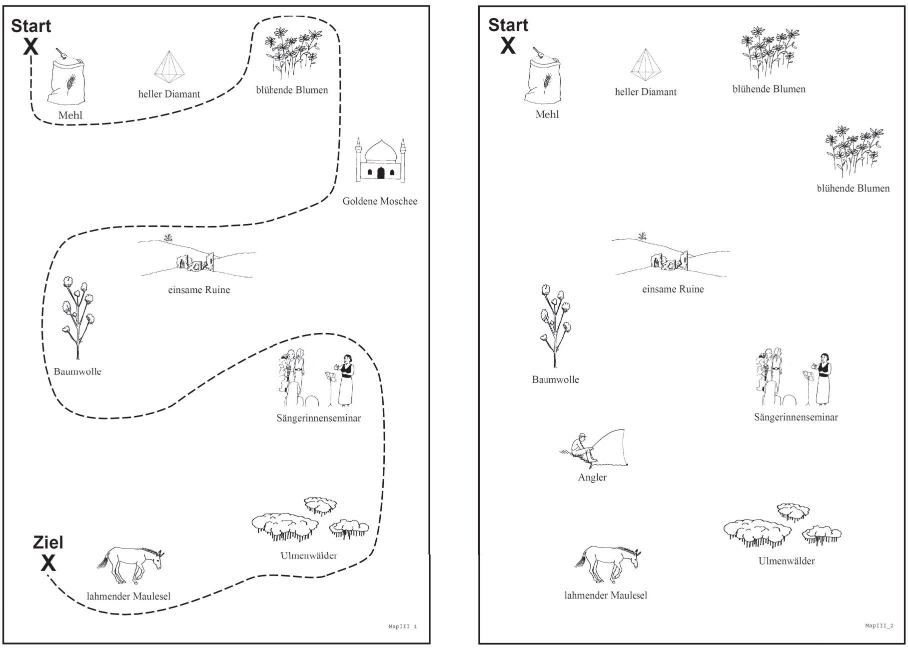
3. *Application*

As a test case for this procedure designed to characterise different styles of intonation, we return to one of the issues mentioned in the introduction: the speech of individuals diagnosed with autism spectrum disorders (ASD). This strikes us as a particularly good such test case due to the contradictory claims in the literature about speakers with ASD as having either robotic or singsong intonation. Although these claims have in part been made several decades ago, the issue has not been resolved in any way since. In the following section, we hope to shed some light on why this might be the case and to demonstrate why our new approach to the characterisation of intonation styles can be helpful in this and other cases.

3.1 Subjects and materials

As part of an ongoing collaboration with the psychiatry department of the University Hospital of Cologne (see e.g. Krüger, Cangemi, Vogeley & Grice, 2018), we have been collecting Map Task recordings (Anderson et al., 1991) between dyads of subjects diagnosed with ASD and dyads of neurotypical (NT) control speakers (all native speakers of German). The materials used in the task are shown in Fig. 8. For the present purpose, we will evaluate data from one female ASD dyad (subjects aged 25 and 46) and one female NT dyad (subjects aged 23 and 26). For each speaker, we extracted 20 excerpts with an average length of 2.5 seconds for further analysis.

Figure 8 - *Map Task materials from the production task. The map for the instruction giver is on the left, the map for the instruction follower is on the right*



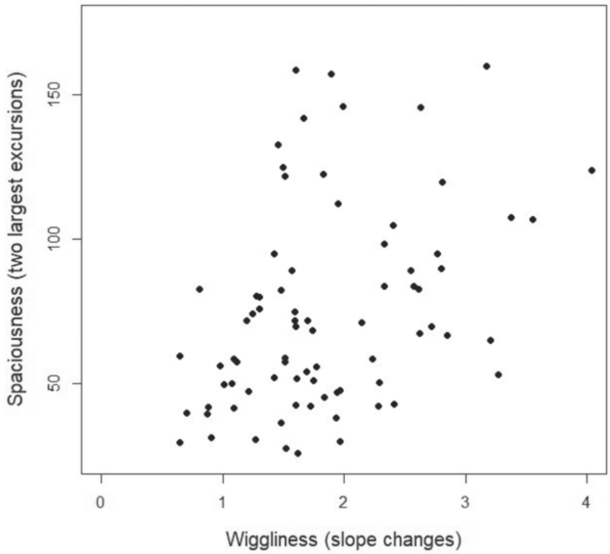
3.2 Results

The results are plotted in Fig. 9. The measure of Wiggliness (Slope Changes) is plotted along the x-axis, the measure of Spaciousness (Maximum Excursions) is plotted along the y-axis.

From a general point of view, the plot seems to show that there is some amount of correlation between the two measures. This is not entirely surprising in itself, but will have to be tested with more data in order to quantify the exact strength of the correlation. For the time being, the pattern nevertheless appears to be clearly in line with the assumption that the two dimensions we have chosen are at least partly independent from each other.

Although the plot in Fig. 9 contains data from two speakers each for the ASD group and for the NT group, it is evident that the data does not cluster into two distinct parts, as would be the case if ASD speakers' intonation was simply either clearly more robotic or clearly more singsongy than that of NT speakers. To make sense of the data, we therefore need to investigate the data at the level of individual speakers.

Figure 9 - *Aggregated data from all 80 excerpts of all 4 speakers. Wigglines is on the x-axis, Spaciousness is on the y-axis*



In Fig. 10 datapoints are colour-coded by speaker. The two ASD speakers are represented by blue and cyan dots, while the two NT speakers are represented by red and orange dots. The NT speaker in red is shown to have a wide range of Wiggliness values and a slightly more limited range of Spaciousness values. This shows that there is a lot of variability in different (parts of) utterances for this speaker. The other NT speaker (in orange) seems to have an intonation style somewhat more towards what could be described as the robotic end. Most values are concentrated in the bottom left quadrant of the plot, representing lower values for both Spaciousness and Wiggliness. Nevertheless, there is variability here, too, with some data points gradually spaced out towards the higher end of both the Wiggliness and Spaciousness scales.

Two examples for the pitch contours represented by the datapoints in Fig. 10 are given in Fig. 11.

Figure 10 - *Wiggliness and Spaciousness values for two NT speakers (red and orange) and two ASD speakers (blue and cyan). The circles in the top right and bottom left of the graph mark the pitch contours shown as examples in Figure 11*

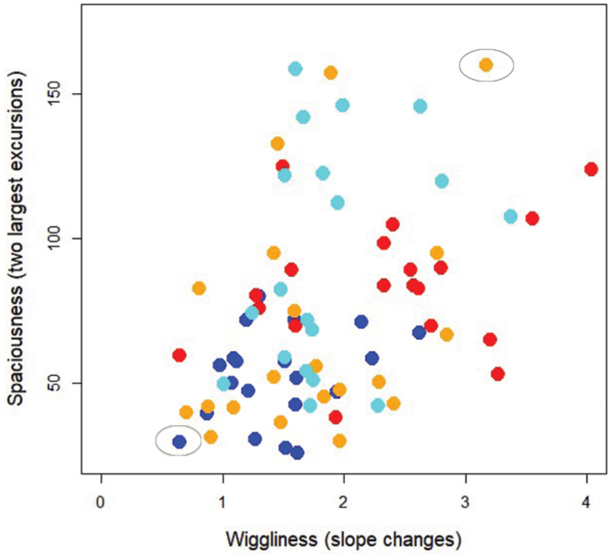
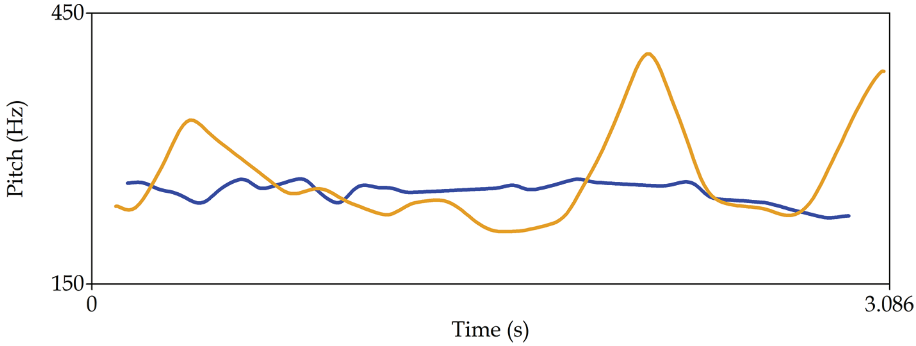


Figure 11 - *Examples of pitch contours represented by dots in Figure 10: The contour in orange is the one marked by a circle in the top right of Fig. 10, the contour in blue is the one in the bottom left of Fig. 10*



Considering the two ASD speakers, the productions of the speaker represented with the blue dots are similar to those of the NT speaker in orange, in being concentrated in the lower regions of both Wiggliness and Spaciousness. The crucial difference between the ASD speaker and the NT speaker is that the productions of the ASD speaker in blue seem to be less variable and therefore much closer to a *uniformly* robotic intonation style, with very few values towards the higher end of Wiggliness and none in the top half of the Spaciousness scale. The second ASD speaker (in cyan) produces a different pattern altogether. Values are spread out along the full range of both dimensions. However, the values are not evenly spread out. There are very few values in the middle of the graph, around the midpoints of Wiggliness and Spaciousness. Instead, values almost seem to be split into one singsong half and one robotic half. The bottom half overlaps with the rather more robotic productions of the ASD speaker in blue, while the top half, taken on its own, might be considered as a typical representation of a singsong style.

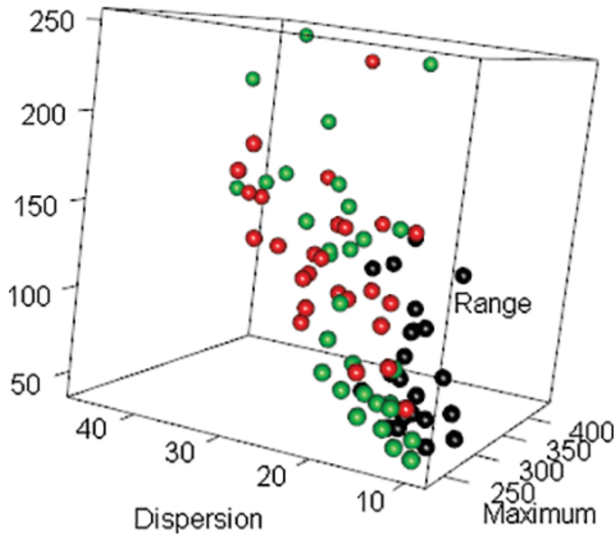
Taken together, the broadest and at the same time most urgent message to be taken from this analysis is that it confirms the absence of a hard dividing line between subjects with a diagnosis of ASD and those without such a diagnosis. Just as autism spectrum disorders *within* themselves cover a range of phenotypical expressions of atypicality that range from low-functioning to high-functioning (amongst other things), there is an *overlap between* the portion of the general population with more autistic-like traits and the portion of the ASD-diagnosed population with outwardly less conspicuous expressions of ASD. This holds true both for general behaviour and for the specific data on intonation styles presented here.

4. *Discussion*

In this contribution, we have pointed out that intonation styles are important in a variety of ways in a number of areas of linguistic inquiry, from the applied to the theoretical. Despite this, accurate and reliable methods for measurement and analysis of what lies behind descriptions of singsong and robotic intonation styles (and what lies in between) have been lacking to date. We have demonstrated the application of a novel, largely automated procedure that fills this gap by reliably quantifying intonation styles, using the example of data from speakers with ASD. These data have also highlighted the necessity of taking into account speaker-specific strategies in the analysis of intonation styles.

Due to the presence of massive individual variability and the absence of clear differences between groups, it is impossible for us to run a conclusive comparison between the metrics employed in this paper and the Long Term Distributional metrics employed in previous research. However, to support the claim that Wiggliness and Spaciousness do indeed provide a better characterisation of speakers' productions, we have plotted utterances from three of the speakers in our dataset into a multidimensional space. The cube in Fig. 12 shows utterance as datapoints, colour-coded per speaker. Points are scattered along the main dimensions of LTD metrics, notably F0 maximum, F0 range and F0 dispersion (calculated as standard deviation of F0 over each individual utterance). The plot indicates that the three LTD metrics are highly correlated, and that they only allow the separation of speakers on the basis of physiological characteristics: Having higher F0 maxima also entails larger values for F0 range and F0 dispersion.

Figure 12 - *LTD measures for three of the speakers in the dataset. F0 dispersion is plotted on the x-axis, F0 range on the y-axis and F0 maximum on the z-axis*



With the approach demonstrated in this contribution we have shown that the picture that emerges from an analysis that is indeed able to capture different dimensions of intonation styles, whilst at the same time giving appropriate consideration to individual differences, confirms the impression that it is inaccurate to describe speakers with ASD as having one particular intonation style. Instead, these speakers seem to show behaviour that goes more towards either end of the spectrum lying between the two poles of "singsongy" and "robotic", both within and across individuals. This reflects similar recent results regarding the prosodic encoding of givenness in ASD (Krüger et al., 2018). Furthermore, our analysis demonstrates that the simplistic labels previously used to describe intonation styles in ASD do not in themselves stand up to thorough investigation.

Although an understanding of the true nature of the data at hand cannot be gained without giving due consideration to individual variability, we submit that this is not the case merely because we are dealing with the somewhat elusive topic of intonation styles in conjunction with the somewhat broad range within the ASD spectrum. In fact, this particular example serves as a useful illustration for understanding the nature and the import of individual variability more generally. Moreover, across different domains of language and different fields of scientific endeavour, not to give due consideration to individual-specific differences is to allow ourselves to be misled by an only apparent simplicity of explanation.

Acknowledgements

We would like to thank Prof. Kai Vogeley of University Hospital Cologne, whose support and expertise were essential to the research presented here. We would also like to thank Harriet Hanekamp and Anika Müller for their invaluable help with data processing. Further, we would like to thank two anonymous reviewers for their extremely helpful comments, as well as the participants of AISV 2018 in Bolzano for insightful and stimulating discussion.

The research for this paper has been funded by the German Research Foundation (DFG) as part of the SFB 1252 "Prominence in Language" in project A02 "Individual behaviour in encoding and decoding prosodic prominence" at the University of Cologne. The first author is partly funded through a doctoral scholarship by the German Academic Scholarship Foundation (Studienstiftung des deutschen Volkes).

Bibliography

Anderson, A.H., Bader, M., Bard, E.G., Boyle, E.H., Doherty, G.M., Garrod, S.C., (...) & Weinert, R. (1991). The HCRC map task corpus. In *Language and speech*, 34(4), 351-366.

Baltaxe, C. (1984). Use of contrastive stress in normal, aphasic, and autistic children. In *Journal of Speech, Language, and Hearing Research*, 27(1), 97-105.

Boersma, P., Weenink, D. (2018). *Praat: doing phonetics by computer* [computer program]. Version, 6.0.39, retrieved 7 April 2018 from http://www.praat.org/.

Cangemi, F. (2015). *mausmooth* [computer program]. Version 1.0, retrieved 7 April 2018 from http://phonetik.phil-fak.uni-koeln.de/fcangemi.html.

Cangemi, F., Grice, M. (2016). The importance of a distributional approach to categoriality in autosegmental-metrical accounts of intonation. In *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1)9, 1-20.

Cangemi, F., Krüger, M. & Grice, M. (2015). Listener-specific perception of speaker-specific production in intonation. In Fuchs, S., Pape, D., Petrone, C., Perrier, P. (Eds.), *Individual Differences in Speech Production and Perception*. Frankfurt am Main: Peter Lang, 123-145.

Celce-Murcia, M., Brinton, D. & Goodwin, J. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.

Coquillon, A.-L., Durand, J. (2010). Le français méridional: éléments de synthèse. In Detey, S., Durand, J., Laks, B. & Lyche, C. (Eds.), *Les variétés du français parlé dans l'espace francophone: ressources pour l'enseignement*. Paris: Ophrys, 243-262.

D'imperio, M., Grice, M. & Cangemi, F. (Eds.) (2016). Advancing prosodic transcription. Special collection in *Journal of the Association for Laboratory Phonology*.

Frota, S. (2016). Surface and structure: transcribing intonation within and across languages. In *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1)7, 1-19.

Graham, C. (2014). Fundamental frequency range in Japanese and English: The case of simultaneous bilinguals. In *Phonetica*, 71(4), 271-295.

Green, H. & Tobin, Y. (2009). Prosodic analysis is difficult… but worth it: A study in high functioning autism. In *International Journal of Speech-Language Pathology*, 11(4), 308-315.

Hirst, D., Di Cristo, A. (Eds.) (1998). *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press.

Hodges-Simeon, C., Gaulin, S. & Puts, D. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. In *Human Nature*, 21(4), 406-427.

Holmes, J. (2013). *An Introduction to Sociolinguistics*, 4th edition. London-New York: Routledge.

Hualde, J., Prieto, P. (2016). Towards an International Prosodic Alphabet (IPrA). In *Journal of the Association for Laboratory Phonology*, 7(1), 5.

Kanner, L. (1943). Autistic disturbances of affective contact. In *Nervous Child*, 2(3), 217-250.

Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). In *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979-1000.

Kuiper, L. (1999). Variation and the norm: parisian Perceptions of regional French. In Preston, D. (Ed.), *Handbook of Perceptual Dialectology Vol. I,* Amsterdam-Philadelphia: John Benjamins, 243-262.

Krüger, M., Cangemi, F., Vogeley, K. & Grice, M. (2018). Prosodic marking of information status in adults with Autism Spectrum Disorders. In *Proceedings of Speech Prosody*. Poznan, 182-186.

Ladd, D.R., Terken, J. (1995). Modelling intra- and inter-speaker pitch range variation. In *Proceedings of XIIIth International Congress of Phonetic Sciences*. Stockholm, 386-389.

Ladd, D.R., Silverman, K.E., Tolkmitt, F., Bergmann, G. & Scherer, K.R. (1985). Evidence for the independent function of intonation contour type, voice quality, and F0 range in signalling speaker affect. In *The Journal of the Acoustical Society of America,* 78(2), 435-444.

Lehiste, I. (1975). The phonetic structure of paragraphs. In Cohen, A., Noteboom, S. (Eds.), *Structure and Process in Speech Perception*. Berlin, Springer-Verlag, 195-206.

Leongómez, J.D., Binter, J., Kubicová, L., Stolarová, P., Klapilová, K., Havlíček, J. & Roberts, C. (2014). Vocal modulation during courtship increases proceptivity even in naive listeners. In *Evolution and Human Behavior*, 35, 489-496.

Liu, H.-M., Kuhl, P.K. & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. In *Developmental Science*, 6, F1-F10.

Mennen, I., Schaeffler, F. & Docherty, G. (2012). Cross-language differences in F0 range: a comparative study of English and German. In *Journal of the Acoustical Society of America*, 131 (3), 2249-2260.

Mohammadi, G., Origlia, A., Filippone, M. & Vinciarelli, A. (2012). From speech to personality: Mapping voice quality and intonation into personality differences. In *Proceedings of the 20th ACM international conference on Multimedia*, 789-792.

Nolan, F. (2006). Intonation. In Aarts, B., Mcmahon, A. (Eds.), *Handbook of English Linguistics*. Oxford: Blackwell, 433-457.

Patterson, D.J. (2000). *A linguistic approach to pitch range modelling.* PhD dissertation, University of Edinburgh.

Paul, H. (1880). *Principien der Sprachgeschichte.* Tübingen: Niemeyer.

Simmons, J.Q., Baltaxe, C. (1975). Language patterns of adolescent autistics. In *Journal of Autism and Childhood Schizophrenia*, 5(4), 333-351.

Torreira, F., Grice, M. (2018). Melodic constructions in Spanish: Metrical structure determines the association properties of intonational tones. In *Journal of the International Phonetic Association*, 48(1), 9-32.

Urbani, M. (2013). *The pitch range of Italians and Americans. A comparative study.* PhD dissertation, University of Padova.