A pitch accent for contrastive emphasis in Danish?

Nicolai Pharao[1], *Rasmus Puggaard-Rode*[2,3]

[1]Department of Nordic Studies and Linguistics, University of Copenhagen, Denmark
[2]Institute for Phonetics and Speech Processing, LMU Munich, Germany
[3]Faculty of Linguistics, Philology and Phonetics, University of Oxford, United Kingdom
nicolaip@hum.ku.dk | rasmus.puggaard-rode@ling-phil.ox.ac.uk

Abstract

This paper presents an exploratory study of a previously undescribed tonal contour in Danish questions. We analyze pitch trajectory data collected for a different purpose using functional principal component analysis and hierarchical clustering, which converge to show that two distinct pitch patterns are used for cueing pragmatic focus in utterance medial position: the 'expected' tonal contour that is generally used to signal stress in Danish regardless of the presence of focus (low–high), and an entirely differently shaped contour (high–low), which we interpret as cueing 'extra' contrastive emphasis. We are (somewhat controversially) proposing that the pattern may constitute a new pitch accent for contrastive emphasis in Danish, although we also call for more targeted research into the intonation of contrastive emphasis in contemporary standard Danish.
Index Terms: intonation, contrastive emphasis, Danish, contour clustering, functional data analysis

1. Introduction

In the standard model of (Standard Copenhagen) Danish intonation, stressed words of two syllables or more are described as having a consistent low–high pitch pattern with a steep rise to the first post-tonic syllable, regardless of modality or where the stressed word appears in a sentence [1]. This model builds on extensive research of read speech by Grønnum (see e.g. [1, 2] for overviews), and has also been shown to hold for spontaneous speech [3, 4].

We recently collected data for a study on co-intrinsic pitch effects in Danish where participants read aloud alternative questions with focused nouns in both utterance-medial and utterance-final position. The pitch trajectories of focused nouns are largely in agreement with the standard model, but listening to the recordings also gave the impression that some speakers placed extra emphatic stress on some utterance medial items. In this paper, we test whether this auditory impression is related to differences in fundamental frequency (F0) contour shape.

In order to explore whether some words in our corpus deviate from the standard model of Danish intonation, we employ two different unsupervised learning techniques, *viz.* functional principal component analysis [5] and hierarchical clustering [6] to test whether there are potentially qualitatively different pitch trajectories in the utterance medial and utterance final words in the data, which would be contra to the standard model.

Our results show that the items which stood out auditorily do indeed have a different pitch pattern, namely a high–low contour with a relatively shallow trajectory. Since the shape of this contour is entirely different from the tonal contour otherwise used to signal stress, we tentatively propose that the pattern constitutes a new pitch accent used to cue contrastive emphasis in Danish. As will become clear, the suggestion is in need of more systematic follow up studies, but we believe that the findings constitute a useful addition to the model of Danish intonation.

2. Danish intonation

The basic component in the model of Danish intonation (developed over an extensive series of studies of read speech and summarized in [1, 2]) is the tonal stress group pattern, which is defined as the pitch contour that is initiated by the vowel of an accented syllable and includes all following unaccented syllables up to the next accented syllable in the utterance. In standard Copenhagen Danish, this tonal stress group pattern is characterized by beginning with a relatively low tone on the accented syllable followed by a steep rise to the first post-tonic syllable and a gradual fall through any following unaccented syllables. This is the pattern found in neutral readings without emphasis. According to the model [1], contrastive emphasis in Danish is signalled by

- more extensive F0 contours associated with the stress group in emphasised words
- often a reduction of the contour in surrounding stress groups in the same utterance
- otherwise qualitatively the same contour in emphasised and non-emphasised words

Importantly, in this model the contour associated with emphasised words is qualitatively the same as that found for non-emphasised words, with the difference limited to the pitch level and magnitude of the excursion. This is illustrated in Figure 1, reproduced from [7], the original study of emphasis in standard Copenhagen Danish. The figure shows stylized pitch contours based on average F0 of each syllable in 6 repetitions of either statements (the panels on the left hand side) or questions (the panels on the right hand side). Each dot represents a syllable and larger dots indicate accented syllables. A triangle indicates the accented syllable in a stress group placed in contrastive emphasis. As can be seen from the figure, all stress groups have the same tonal contour: a low tone followed by a rise and then followed by a fall (the exception being perhaps the reduced stress group patterns in the utterances where one is emphasized; here the rise to the first post-tonic is considerably less extensive and in some cases non-existent).

This qualitatively invariant stress group pattern, regardless of level of emphasis, is in contrast to other languages where contrastive emphasis or focusing may be signalled by a pitch accent with a different contour [8]. The model for Danish is based on read speech, but [3, 4] have empirically verified the qualitative invariance of the stress group pattern in spontaneous speech, [3] also for words with contrastive emphasis.

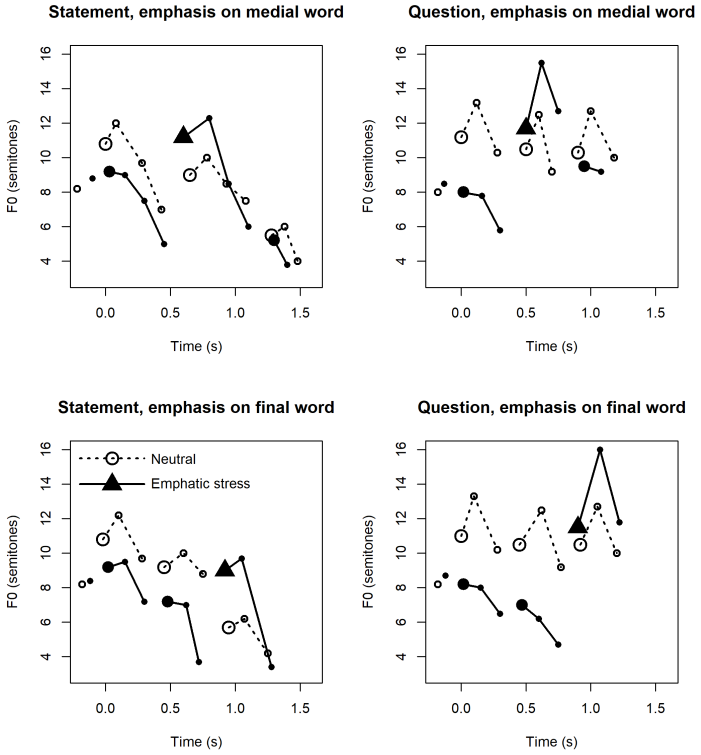

Figure 1: *Stylised stress group patterns in Copenhagen Danish with and without emphatic stress. Reproduced from [7].*

3. Methods and materials

This exploratory study is conducted on read aloud sentences contrasting two phrases in an interrogative frame (based on the design from [9]) and collected for different purposes. 12 native speakers of Danish from Greater Copenhagen each read aloud 132 constructed alternative question sentences following the mould 'Did you verb X or Y?', where each target word occurred in both X and Y positions. An example would be *Er det **dine** eller er det **mine**?* 'Are they yours or are they mine?', designed to elicit high global pitch on the first focused item (*dine*) and low global pitch on the second (*mine*). These utterances invite the speaker to contrast the word in utterance medial position (X) with the word in utterance final position (Y). While such a construction could potentially elicit a higher degree of prominence on the words in utterance-medial position, due to their being pragmatically focused, there should be no difference in the shape of the fundamental frequency contour for these items, according to the standard model model [1].

Sentences were prompted using SpeechRecorder [10] in a sound-attenuated booth at the University of Copenhagen. The raw data are freely available via the Open Science Framework (DOI: 10.17605/OSF.IO/67SHC). Pitch was extracted from 790 suitable tokens using the modified harmonic sieve algorithm implemented in the `wrassp` library in R [11] with liberal pitch floor and ceiling of 75 Hz and 500 Hz, respectively. Pitch was extracted at 20 time-normalized steps. F0 was normalized using octave-median scaling by speaker [12].

The data were analysed in R [13] using unsupervised learning algorithms, *viz.* functional principal component analysis (FPCA) as implemented in the `fdapace` library [14], and hierarchical clustering as implemented in the `contour_clustering_gui` [6].

4. Results

4.1. Functional principal component analysis

In order to uncover the principal modes of F0 variation in utterance medial position, we analysed the contours using FPCA (see e.g. [5, 16]). FPCA is used to decompose the multifaceted sources of variation in a range of functions (i.e. curve shapes, such as F0 trajectory shapes) into a small number of continuous variables, corresponding to unique modes of variation. An overall average trajectory shape $\mu(t)$ is estimated, as well as a number of supplementary trajectory shapes (principal components, $PC1 \ldots n$), each of which capture a principal mode of variation in the input data. Each trajectory in the input data receives a score for each PC $s_1 \ldots_n$ indicating how closely it matches this mode of variation, which makes it possible to reconstruct the gross shape of an input trajectory by weighing $\mu(t)$ with that trajectory's PC scores (see [17]).

The analysis of the Danish stress group data shows that 3 PCs account for $>95\%$ of the variance in F0 contour shape. By far the most informative of these is PC1, which accounts for approx. 72% of the variance. Since PC2 and PC3 did not turn out to be very informative, we do not discuss them further here.

PC1 is illustrated in Figure 2. The grey line vertically in the middle of Figure 2 illustrates the average trajectory shape $\mu(t)$. Increasingly red shades illustrate higher values of s_1, while increasingly blue shades illustrate lower values of s_1. (Note that the signs of PC scores are randomly assigned). $\mu(t)$ has the expected shape of a Danish disyllabic stress group, with low pitch on the accented syllable with a steep rise presumably leading

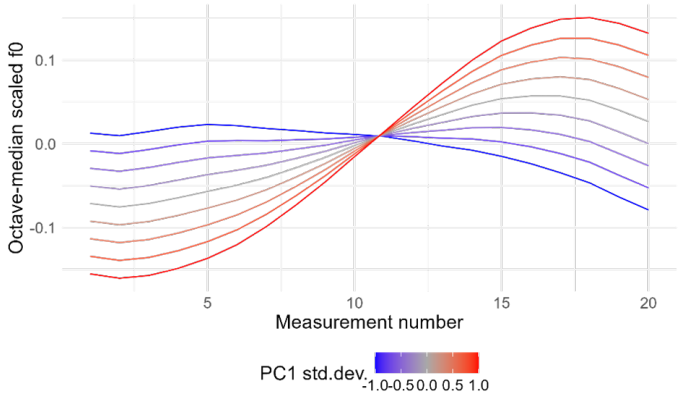

Figure 2: *Illustration of relatively high and low scores of PC1 relative to the average pitch trajectory. Plotted in R using* `ggplot2` *[15]*

into the post-tonic syllable. Items with high s_1 essentially show the same pattern but with more pronounced extrema, i.e. with lower pitch on the accented syllable and a steeper rise and eventually higher pitch on the post-tonic. Items with low s_1, on the other hand, show the reverse of the expected pattern: a relatively high initial F0 on the accented syllable, followed by a shallow dip leading into the post-tonic. To our knowledge, this high–low pattern has not previously come up in acoustically based descriptions of Copenhagen Danish intonation. (Note that this is unlike some other regional varieties of Danish, where this has been described as the default pattern for stressed syllables [18]).

4.2. Hierarchical clustering

Using the contour clustering tool developed by [6], which is specifically designed to assign qualitatively different F0 contours to separate categories, we conducted hierarchical clustering of the F0 trajectories in target words occurring in either medial or final position. 308 contours from utterance-medial tokens were eligible for clustering analysis. After removing 87 tokens with a level contour (likely due to the inconsistent pitch tracking and creaky voice typical of Danish *stød* [19, 20, 21]), 221 contours in medial position were subjected to clustering analysis. Stepping down from 25 possible clusters (following [6]), the final analysis revealed only two distinct contours, shown in Figure 3.

Similar to what we saw above in the FPCA results, the predominant contour is a rise from a low tone in the accented syllable (shown in the left panel), but the opposite pattern is also found (albeit only in a relatively small number of tokens), i.e. a high tone on the accented syllable followed by a fall to the post-tonic. This high–low pattern is contra to the model in [1] and the findings in [3].

Analysing the same words in final position yielded 193 contours available for cluster analysis, after removal of flattened contours due to *stød* (as well as eleven spurious contours) to ensure comparability. The resulting two clusters, shown in Figure 4, reveal a slight difference in level of the contours only, i.e. there is no difference in contour shape. As such, at least in this data, the high–low contour is only found in utterance medial position, which is arguably more likely to be parsed as contrastive emphasis in the sentence format that we analyse here.

When we compare the results of the clustering analysis in utterance-medial position to the FPCA results reported above, we find near-perfect correspondence between the information captured by PC1 and that captured by the two clusters. As shown in the left panel of Figure 5, items in cluster 1 (with the expected low–high pitch trajectory) generally have a relatively high s_1, whereas items in cluster 2 (with the unexpected high–low pitch trajectory) generally have quite low s_1, to the point where there is next to no overlap in s_1 between the two clusters: items with s_1 <-0.25 are almost certainly in cluster 2. The right panel in Figure 5 shows the s_1 and s_2 values of each item coloured by which cluster the item was assigned to, and again we see a relatively small number of items in cluster 2 which consistently have low s_1. s_2, which we otherwise do not discuss here, is not related to the clustering outcome.

The fact that FPCA and hierarchical clustering both support the existence of two different contours with a remarkable consistency suggests that while the high–low pattern may be comparatively rare, it is a robust pattern in the data.

5. Discussion

While the analysis presented here suggests the option of a specific contour shape for contrastively emphasised words, it is important to note that only 6.7% (n = 33) of the analysable contours were realized with this high–low pattern. Contours included in the high–low cluster were only attested for 9 out of the 12 speakers, and none of these speakers exclusively used the high–low pattern for emphasised words in medial position. In other words, the attested high–low pattern is at most an optional way to mark emphasis prosodically in contemporary standard (Copenhagen) Danish, albeit one that has not previously been attested. The high–low stress group pattern is not unknown in Copenhagen Danish, however. In fact, this pattern was the standard cue for accented syllables in Copenhagen Danish up until the 1920s [22], and informal observations suggest that the pattern may still occur in highly formal settings. It does not seem unreasonable that a pitch contour that is today perceived as highly conservative could be used for marking extra emphasis in focus positions, given the ideological link between conservative speech and distinct pronunciation in Danish [23]. In other words, the high–low contour may have previously functioned as a generic signal for stress, but when this becomes a conservative and increasingly obsolete way of signalling stress, the contour was freed up to take on the function of emphasis [24].

This study has only been a first step towards describing the use of high–low pitch contours as a potential cue to contrastive emphasis in Danish. The study relies on data not col-

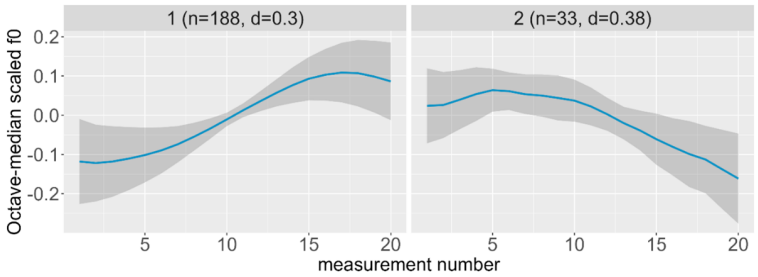

Figure 3: *Clusters in utterance medial position, plotted using* `contour_clustering_gui` *[6].*

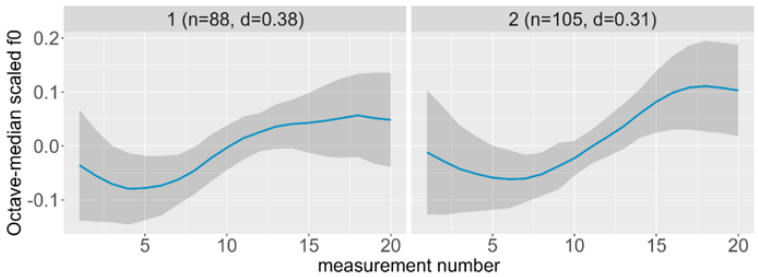

Figure 4: *Clusters in utterance final position, plotted using* `contour_clustering_gui` *[6].*

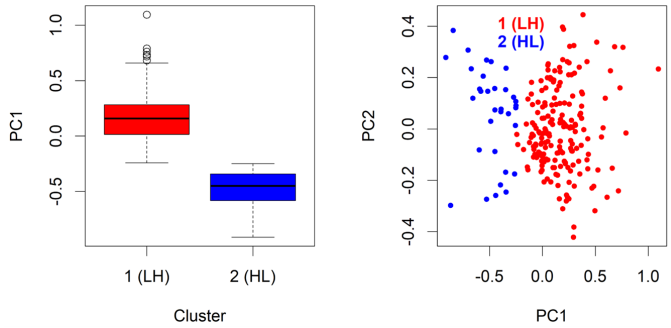

Figure 5: *Comparison of utterance-medial clusters and PC scores.*

lected specifically for this purpose, and while the utterance medial position in an alternative question sentence arguably can elicit contrastive emphasis, it is only relatively certain to elicit focus. This may have led to a certain underestimation of how often the high–low pattern is actually used to cue contrastive emphasis, and how many speakers have it as part of their intonational 'toolbox'. In order to understand the pattern more fully, more targeted research is required, specifically studies that are designed to elicit contrastive emphasis, to see if the high–low pattern is potentially a tool that (more) speakers reach for more often than uncovered here, and to test whether there are other restrictions (e.g. semantic) that we have not uncovered here.

In such further explorations, it would be interesting to study whether there might be a trade-off between the shape of the contour on the stress group with contrastive emphasis and the shape and particularly the pitch range of the contour associated with neighbouring stress groups. In the present study, we have only looked at the focussed stress groups themselves, but recall from Grønnum's 1980 study introduced [7] above that contrastive emphasis may also affect neighbouring stress groups in Copenhagen Danish, at least in read speech. Compared to segmentally and lexically identical utterances without emphatic accentuation, the stress groups surrounding the emphatically accented stress group have a diminished pitch range, in particular with respect to the extent of the rise from the accented to the first posttonic syllable, which sometimes even appears to be absent. This modification is seen as evidence of the relative nature of prominence: the reduction of the surrounding stress groups makes the emphatic stress stand out even more, acoustically and perceptually. This holds for utterances where the emphatic stress is cued solely by a more extensive pitch range in the focussed stress group, i.e. where the low-high pattern is retained. It remains to be seen whether speakers, when they use the high–low pattern documented here, continue to compress the tonal range of the surrounding stress groups, or whether they resemble the range found in utterances without any emphatic stress. One may speculate that when the high–low pattern is employed, there is less of a need to modify the surrounding stress groups as when the low–high pattern is retained. In other words, when the shape of the contour is itself remarkable, it may function more independently as a local pitch accent proper, irrespective of the surrounding stress groups.

6. References

[1] N. Grønnum, "Modeling Danish intonation," in *Prosodic Theory and Practice*, J. Barnes and S. Shattuck-Hufnagel, Eds. Cambridge & New York: The MIT Press, 2022, pp. 85–119.

[2] N. Grønnum, *The groundworks of Danish intonation. An introduction*. Copenhagen: Museum Tusculanum, 1992.

[3] N.-J. Dyhr, "The fundamental frequency in Danish spontaneous speech with special references to syllables boosted for emphasis," in *Aspects of Danish Prosody*, J. Rischel and H. Basbøll, Eds. Odense Universitetsforlag, 1995, pp. 49–67.

[4] J. Tøndering, "Intonation og informationsstatus. Belyst ved prominensmålinger," Unpublished MA Thesis, University of Copenhagen, 2004.

[5] M. Gubian, F. Torreira, and L. Boves, "Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts," *Journal of Phonetics*, vol. 49, pp. 16–40, 2015.

[6] C. Kaland, "Contour clustering. A field-data-driven approach for documenting and analysing prototypical f0 contours," *Journal of the International Phonetic Association*, vol. 53, no. 1, pp. 159–188, 2023.

[7] N. Thorsen, "Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish," *Annual Report of the Institute of Phonetics, University of Copenhagen*, vol. 14, pp. 121–205, 1980.

[8] C. Gussenhoven, *The Phonology of Tone and Intonation*, ser. Research Surveys in Linguistics. Cambridge: Cambridge University Press, 2004.

[9] J. Kirby and D. R. Ladd, "Effects of obstruent voicing on vowel F0. Evidence from "true voicing" languages," *Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 2400–2411, 2016.

[10] C. Draxler and K. Jänsch, "SpeechRecorder. A universal platform independent multi-channel audio recording software," in *Proceedings of the Fourth International Conference on Language Resources and Evaluation*, M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa, and R. Silva, Eds. Lisbon: European Language Resources Association, 2004, pp. 559–562.

[11] R. Winkelmann, J. Harrington, and K. Jänsch, "EMU-SDMS. Advanced speech database management and analysis in R," *Computer Speech & Language*, vol. 45, pp. 392–410, 2017.

[12] C. De Looze and D. Hirst, "The OMe (Octave-Median) scale: A natural scale for speech melody." in *Proceedings of Speech Prosody*, N. Campbell, D. Gibbon, and D. Hirst, Eds., vol. 7. Dublin: International Speech Communication Association, 2014, pp. 910–914.

[13] R Core Team, "R. A language and environment for statistical computing," 2025, v. 4.5.0.

[14] Y. Zhou, S. Bhattacharjee, C. Carroll, Y. Chen, X. Dai, J. Fan, A. Gajardo, P. Z. Hadjipantelis, K. Han, C. Zhu, H.-G. Müller, and J.-L. Wang, "`fdapace`. Functional data analysis and empirical dynamics," 2022, R package version 0.5.9.

[15] H. Wickham, *ggplot2. Elegant graphics for data analysis*, ser. Use R. New York: Springer, 2016.

[16] J. O. Ramsay and B. W. Silverman, *Functional data analysis*, 2nd ed., ser. Springer Series in Statistics. New York: Springer, 2005.

[17] R. Puggaard-Rode, *Stop! Hey, what's that sound? The representation and realization of Danish stops*, ser. LOT Dissertation Series. Amsterdam: Netherlands Graduate School of Linguistics, 2023, no. 631.

[18] N. Grønnum, "Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German," *Phonetica*, vol. 47, no. 3-4, pp. 182–214, 1990.

[19] E. Fischer-Jørgensen, "Phonetic analysis of the stød in Standard Danish," *Phonetica*, vol. 46, no. 1/2/3, pp. 1–59, 1989.

[20] N. Grønnum, "Three quarters of a century of phonetic research on common Danish stød," *Nordic Journal of Linguistics*, vol. 46, no. 3, pp. 299–330, 2023.

[21] M. Watkins, R. Puggaard-Rode, P. Boersma, and S. Hamann, "Robustness of F0 Ratio as a diagnostic. Comparing creaky voice in Danish and Seoul Korean," in *Proceedings of Interspeech 2025*, O. Scharenborg, C. Oertel, and K. Truong, Eds. Rotterdam: International Speech Communication Association, 2025, pp. 4203–4207.

[22] L. Brink, "Udtale: Stød, tryk og tonegang," in *Dansk sproghistorie 2. Ord for ord for ord*, E. Hjorth, Ed. Aarhus: Aarhus Universitetsforlag, 2018, pp. 229–254.

[23] T. Kristiansen, "Language attitudes and language politics in Denmark," *International Journal of the Sociology of Language*, vol. 159, pp. 57–71, 2003.

[24] P. Eckert, "Three waves of variation study. The emergence of meaning in the study of sociolinguistic variation," *Annual Review of Anthropology*, vol. 41, pp. 87–100, 2012.