

MULTIPLE PROSODIC PARAMETERS SIGNALING INFORMATION STRUCTURE: PARALLEL FOCUS MARKING IN FINNISH

Anja Arnhold

Goethe University Frankfurt am Main, Germany; University of Helsinki, Finland
anja.arnhold@gmail.com

ABSTRACT

The present production study shows that information structure is prosodically expressed in several parallel ways in Finnish. In addition to producing words in narrow focus with (1) a larger duration, (2) higher f_0 peaks and greater falls, and (3) an earlier and more consistent peak alignment, participants frequently marked these words using also (4) pauses and (5) non-modal voice quality. Furthermore, the means of prosodic marking of information structure interact with sentence position in shaping the prosodic form of the word.

Keywords: prosody, focus marking, Finnish

1. INTRODUCTION

Prosodic marking of focus in Finnish has almost exclusively been studied in terms of pitch scaling and duration. Researchers usually agree that an expanded pitch range on the rise-fall word accents signals focus (e.g. [10]). More specifically, [9] found increased rises on medial, directly pre-final, words and extended pitch falls in final position. In turn, the durational marking of focus is restricted to contrastive accents according to [8].

This paper presents a production study assessing and extending the previous studies by considering additional strategies of prosodic focus marking. Moreover, the study was designed to take into account two further factors that can be expected to influence the prosodic form of Finnish words, namely the lexically distinctive vowel quantity and the position of the word in the sentence. The preliminary analyses presented here will only consider the interaction of the factors focus and position. As will be shown, this will not only allow us to assess previous claims and create novel hypotheses about the general prosodic structure of Finnish, but it will also offer new insight into the theory of prosodic marking of information structure in general.

2. METHODS

2.1. Materials

Eight three-word sentences in unmarked SVO word order were embedded in seven different information structural contexts: with broad focus (all-new sentence), with narrow information focus on the subject, verb and object, respectively, and with narrow corrective focus on each of the three constituents, respectively. When one word was narrowly focused, the other two were not only unfocused, but also contextually given. All words were disyllabic, with four sentences containing only long and the other four only short vowels, as far as this is allowed by the restrictions of the Finnish case marking morphology. Note that the position of a word in the sentence was correlated with its grammatical function. This correlation would be difficult to avoid as variations in word order are tightly connected with information structure in Finnish [11].

2.2. Participants and recordings

Seventeen native Finnish speakers (10 female) from Helsinki or the surrounding area participated in the study. Data from an additional participant was not considered in the analysis, because the tonal contours were very flat as a whole. The participants uttered the sentences as answers to pre-recorded questions inducing the designated information structure. The 56 target sentences were interspersed with 110 filler sentences, using a different pseudo-randomized order for each participant. The materials were recorded in a sound-treated room, using a high-quality condenser microphone and an analogue to digital converter.

2.3. Editing and analysis

One sentence was skipped during the recording process and four other sentences were excluded

due to hesitations or slips of the tongue, so that altogether 947 sentences were included in the analysis. Acoustic measurements were done with Praat [3]. The sounds were segmented and the f_0 turning points identified automatically with additional manual verification. Two annotators, the author and a phonetically trained student with no knowledge of the objective of the study, marked stretches of speech that were articulated with non-modal voice quality for half of the sentences each. Peak alignment was calculated as the distance of the word's f_0 maximum from the end of the first syllable nucleus, following [7]. Maxima differing less than 0.8st from the neighboring minima were excluded, following an observation by [4].

Unless stated otherwise, all statistical analyses were done by fitting linear mixed-effect models as implemented in R to the data. Speakers and lexical items were included as a crossed-random factor in all models. Apart from the binomial models for voice quality and occurrence of pauses, p-values were obtained with Markov chain Monte Carlo simulation [2, 6]. All reported effects were significant at or beyond the $p < 0.05$ level.

3. RESULTS

3.1. Duration

All three constituents were realized with greater average word duration when they were in narrow focus than for the broad focus intercept. Contrary to the claim that duration is used to mark only contrastive accent [8], this was found to be the case also when only information focus was considered. Additionally, unfocused given words were shortened relative to broad focus realizations.

In addition to a main effect of position leading to longer durations later in the sentence, interactions of two kinds also emerged: First, verbs in broad focus were relatively short, so that the difference between given and broad focus was not significant in sentence-medial position and the effect of narrow focus was significantly larger for these words than for words in initial position. Second, the lengthening effect of narrow focus was less strong in final position than for the first position intercept, although still significant.

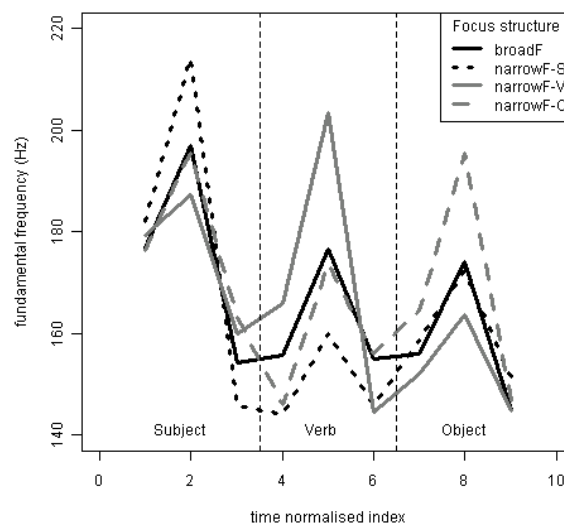
3.2. Pitch scaling

In comparison to broad focus realizations, narrow focus words showed higher f_0 maxima, an effect that was significantly larger in medial and final position than sentence-initially. For given words,

the peak height was reduced, the effect being larger later in the sentence, although the interaction was only significant regarding the final position.

Differing from the findings of [9], f_0 falls were significantly increased in narrow focus for all positions, while the increased magnitude of the rises was only significant in the medial and final position.

Figure 1: Time-normalized sentence contours for different focus structures, displaying the average f_0 maximum and preceding and following f_0 minimum for each word.



3.3. Peak alignment

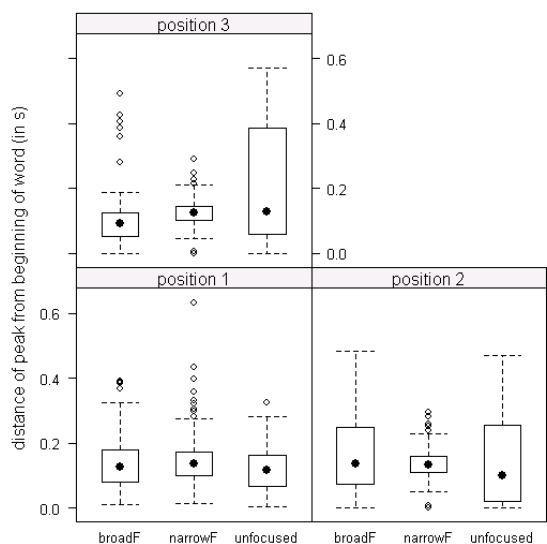
Figure 2 shows a much smaller variation in peak alignment, as measured in the present analysis, in the narrow focus condition for medial and final words (for SDs : $t(16) = -4.81$, $p < .001$ and $t(16) = -4.69$, $p < .001$, respectively). Additionally, f_0 maxima were aligned significantly earlier in narrow focus than in the broad focus intercept. This effect was modulated by an interaction with position: It is significantly larger for sentence-medial than sentence-initial words and not significant for words in final position. A further interaction showed that peaks of unfocused given words were aligned significantly later only in final position. Overall, the peak alignment was earlier for sentence-final words.

3.4. Pauses

Altogether, the speakers produced 111 pauses in the 947 sentences investigated here. Of these, 60 occurred after narrowly focused subjects. In this condition, a pause was realized in 22% of the cases,

whereas a pause occurred after 10% of the narrow focus verbs. The increased use of pauses after a word in narrow focus was statistically significant, as was the difference between sentence-initial subjects and sentence-medial verbs, but no significant interactions emerged.

Figure 2: Boxplot comparing peak alignment for words in broad focus, narrow focus and unfocused given words in initial, medial and final position.

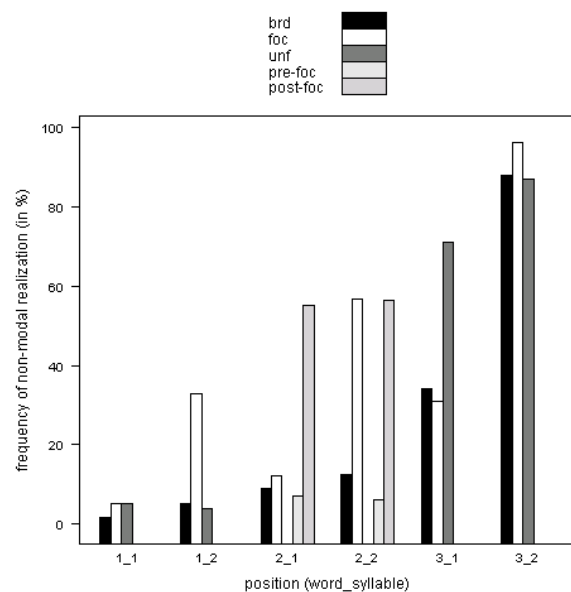


3.5. Voice quality

Figure 3 illustrates the percentage of syllables that were (partly or completely) realized with non-modal voice quality in different focus conditions for each position in the sentence. For example, the bars labeled “2_1” refer to the first syllable of the sentence-medial verbs, the only position in which given words occurred both pre- and post-focally. As Figure 3 shows, this difference was crucial in the usage of non-modal phonation. Therefore, the data was re-modeled using a single four-level factor focus condition (broad focus, narrow focus, pre-focal given and post-focal given). Model comparison (*anova* in R) showed that this gave a significantly better fit than the models with three-levels for focus (broad focus, narrow focus and given) and position. The model for the words’ first syllable indicated that the percentage of non-modal realizations was significantly increased only on post-focal given words. For the second syllable, a significantly higher amount of non-modal realizations emerged for narrowly focused and post-focal given words, while on pre-focal given words, non-modal voice quality appeared less often than in broad focus.

In other words, the speakers often marked narrow focus by using non-modal phonation on the second syllable of the focused word and on the following syllables. Also, the usage of non-modal voice increased later in the sentence, with both syllables of words in medial and final position differing from the first position intercept.

Figure 3: Proportion of non-modal voice quality for syllables in all positions as a function of focus condition.



4. DISCUSSION

The present results illustrated that at least five different prosodic strategies of focus marking are available to speakers of Finnish. Interestingly, speakers simultaneously used several parallel means of conveying the contextually induced focus structures, giving rise to redundancy. It is likely that this multiple focus marking is due to the necessity to transmit other linguistic distinctions at the same time through the same channels. This hypothesis is supported by the numerous interactions between focus condition and position.

For example, the present study identified post-focal pauses as a means of prosodic focus marking, but this strategy was used more often for subjects than verbs. It is suggested here that these pauses are an indication of prosodic phrasing. Although details of this account are left to be clarified in later work, it is indicative that the less frequent post-verbal pauses would introduce a VP-internal boundary. This suggests that prosodic phrases also mark syntactic structure, restricting the utilization of pauses for other purposes.

A similar picture emerged for the marking of finality. Duration and non-modal phonation, which have previously been identified as finality markers [5], were found to signal focus structure in the present data. Thereby, lengthening of narrow focus words was smaller in final position while overall durations were longer, implying that final lengthening restricted the effectiveness of duration as a focus cue in this position. Also the marking of narrow focus through non-modal second syllables is less clear in final position, as Figure 3 shows. Interestingly however, the percentage of non-modal realizations was found to increase over the sentence instead of stepping up on the final word, contrasting with previous findings [8].

Lastly, prosodic marking of predicates affected the expression of focus structure. This is in line with previous research showing that finite verbs are special with respect to f0 excursions: First, unfocused given realizations were more similar to broad focus ones than for the other constituents. Second, narrow focus differed more strongly from broad focus than for the other constituents [1, 10]. Additionally the same pattern was observed here also for duration and peak alignment. However, the account that finite verbs only bear accents in narrow focus [10] does not capture the f0-differences between unfocused and broad focus words (see also [1]). Still, the present results suggest that it might be fruitful to use the term accent to cover a larger range of phenomena and to attribute to it more correlates than f0 and duration.

Therefore, leaving details like accent type and distribution open for the time being, the following hypothesis is put forward: Finnish accents are expressed by f0 excursion and duration [8], but they are not restricted to contrastive accents, and they are additionally signaled by the timing of f0 movements and the use of non-modal voice quality. It is suggested further that narrowly focused words receive the nuclear accent of the sentence, as do the words in sentence-final position (as the focus exponents) in broad focus. This hypothesis would account for the missing distinction in peak alignment for narrow and broad focus in the final position. It could also explain why non-modal phonation is usually found only after the first syllable of the focus exponent, since it is arguably restricted by the necessity to realize the nuclear accent. This preliminary account will be re-assessed by a more detailed analysis of the data and further ongoing experimentation in the near future.

5. CONCLUSION

Whereas previous studies have concentrated on investigating f0 and duration, the present paper showed that speakers of Finnish use a considerably wider range of prosodic means to express information structural variation. Focus is marked by the size of f0 excursions, alignment of f0 peaks, duration, voice quality and post-focal pausing. Additionally, the focus effects were modulated by the position of the word in the sentence.

The outlined account of parallel prosodic marking of focus and its interaction with the prosody of Finnish underlines the need to consider also prosodic parameters beyond f0 and duration as important correlates of information structure.

6. ACKNOWLEDGEMENTS

This work was supported by a grant from the German National Academic Foundation.

7. REFERENCES

- [1] Arnhold, A., Vainio, M., Suni, A., Järviö, J. 2010. Intonation of Finnish verbs. *Proc. 5th Speech Prosody* Chicago, 100054, 1-4.
- [2] Baayen, R.H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- [3] Boersma, P., Weenink, D. 2010. Praat. Doing phonetics by Computer (Version 5.1.29). <http://www.praat.org>
- [4] Fry, D.B. 1958. Experiments in the perception of stress. *Language and Speech* 1(2), 126-152.
- [5] Nakai, S., Kunnari, S., Turk, A., Suomi, K., Ylitalo, R. 2009. Utterance-final lengthening and quantity in Northern Finnish. *J. Phon.* 39(1), 29-45.
- [6] R Core Team. 2010. R: A Language and Environment for Statistical Computing (Version 2.11.1). <http://www.r-project.org>
- [7] Schepman, A., Lickley, R., Ladd, D.R. 2006. Effects of vowel length and right context on the alignment of Dutch nuclear accents. *J. Phon.* 34(1), 1-28.
- [8] Suomi, K., Toivanen, J., Ylitalo, R. 2008. *Finnish Sound Structure. Phonetics, Phonology, Phonotactics and Prosody*. Oulu: University of Oulu.
- [9] Vainio, M., Järviö, J. 2007. Focus in production: Tonal shape, intensity and word order. *J. Acoust. Soc. Am.* 121(2), EL55-EL61.
- [10] Välimaa-Blum, R. 1993. A pitch accent analysis of intonation in Finnish. *Ural-Altaische Jahrbücher* 12, 82-94.
- [11] Vilkuna, M. 1995. Discourse configurationality in Finnish. In Kiss, K.É. (ed.), *Discourse Configurational Languages*. New York/Oxford: Oxford University Press, 244-268.