PHONETIC EFFECTS OF CORRECTIVE FOCUS IN ESTONIAN

Nele Salveste, Jonathan Harrington, Felicitas Kleber

Institute of Phonetics and Speech processing, Ludwig Maximilian University of Munich
nele.salveste@gmail.com, {jmh; kleber}@phonetik.uni-muenchen.de

ABSTRACT

The aim of this study was to investigate whether *narrow corrective focus* causes stronger intonational emphasis than *narrow new information focus.* This research question was tested for Estonian, a free word order language where an object noun phrase at the beginning of the sentence (e.g. OVS word order) is claimed to be in structural focus position. We predicted greater pitch prominence for sentence-initial objects in focus. The acoustic parameters of pitch prominence measured in this study – peak height, slope, peak alignment – did not vary as a function of different focus types and word orders. The results therefore support semantic models (e.g. [19]) by which corrective and new information focus do not differ in terms of acoustic prominence.

Keywords: pitch prominence, narrow focus, corrective focus, free word order, Estonian.

1. INTRODUCTION

In pragmatics, it has been proposed that there are a number of different kinds of narrow focus (see e.g. [8]). Two commonly investigated subtypes are *new information focus* and *correction*. The function of *new information focus* is to provide novel information to the discourse and it arises in the context of *wh*-question or in the context of underlying *wh*-question. C*orrective narrow focus* rejects and replaces the incorrect information from the preceding context. See the examples in (1) and (2) respectively (replicated from [4]).

(1) A: Who fried an omelette yesterday?
B: DAMON fried an omelette yesterday.
(2) A: Did Harry fry an omelette yesterday?
B: DAMON fried an omelette yesterday.

The semantic approach to focus (e.g. [19]), on the other hand, does not differentiate between the two foci in the examples of (1) and (2). Any kind of focused expression is already contrasted to the set of other possible entities that could occupy the same position in the utterance. Thus, being contrastive is the inherent property of focus [19]. According to [19], the only difference between the examples in (1) and (2) would be the size of the set of alternatives: in (1B) the set of alternatives consists of all the entities that can figure as agents, whereas in (2B) the set of alternatives is made previously salient and consists of only one entity: *Harry*. The question is, whether the size of the set of alternatives needs to have some prosodic or phonetic effect.

The pragmatic theory of information structure described in [5] suggests that the narrow corrective focus is intonationally more emphatic than narrow new information focus. Phonetic studies of intonation [1,4,6,9,11,20] report inconsistent results on intonational emphasis of corrective focus. Some studies [6,9,11] have presented evidence for higher scaling of high pitch accent (H*) in corrective focus for English, while [4] found, somewhat contrary to the expectations, lower pitch peaks. [9] demonstrated for Dutch that corrective as opposed to new information focus is characterized by a steeper F0 fall. Other studies [1,20] have found no clear intonational differences between these two types of narrow focus. From perceptual studies of pitch prominence in English and German it is known that peaks that are both higher [12,14,18] and aligned later [12,14] cause perception of stronger emphasis. Based on these results, we predict that peaks in corrective focus are higher and later aligned than peaks in new information focus.

Prosodic means of focus in Estonian are not very well established. Two studies show that narrow focus is signalled by pitch prominence in terms of higher F0 and accent shift [20,21]. However, whether corrective focus is expressed by different intonational cues is still not clear. In addition, different word orders (SVO *vs.* OVS) cue the pragmatic interpretation of a sentence constituent in Estonian. The OVS word order induces emphatic, possibly corrective reading on the object (cf. also [7, 15]). Since it has been shown that both word order and intonation jointly signal focus [21], we predict that word order contributes to the production of additional emphasis on sentence-initial objects. We expect higher and later peaks for object than for subject noun phrases at the beginning of the sentence. In addition, we test whether the different word orders affect the degree of declination: the object noun phrase in OVS word order might attract nuclear prominence which may cause the topline slope of declination to be steeper in a phrase with OVS word order than in a phrase with SVO word order. [8] suggests that the correctiveness might be

signalled on the prenuclear prosody. Thus, if the corrective focus results in stronger pre- or postfocal F0 compression, then the peak difference should vary as a function of focus type as well.

2. EXPERIMENT

The effects of focus type (narrow new information *vs.* narrow corrective focus) and word order (SVO *vs.* OVS) were tested in a speech elicitation task.

2.1. Materials

Four sentences were constructed as *target sentences* which consisted of three two-syllable words with mostly sonorous sounds. The target sentences were in SVO order that were permuted to OVS structures, so that all subjects and objects occurred sentence-initially as well as sentence-finally.

The sentences were triggered as responses to questions or assertions that we will call *context*. See the examples (3) and (4) for the SVO word order. There was another set with OVS word order where focus was elicited similarly.

(3) new information focus
Context 1: Somebody drew a whale?
Context 2: Leena drew something?
Target: *Leena maalis vaala*. (Lena drew a whale.)

(4) corrective focus
Context 1: Anna drew a whale.
Context 2: Leena drew some poppies.
Target: (Ei!) *Leena maalis vaala*. (No!) Leena drew a whale!

The contexts in (3) were *wh*-questions that were signalled by intonation. The presentation of this form of *wh*-question typically involves more detailed sentence responses. The contexts in (4) were designed to elicit corrective focus either on the object or on the subject.

Two word orders (SVO, OVS), two types of foci (new information *vs.* corrective) crossed with two grammatical noun phrases (object *vs.* subject) resulted in 8 conditions. The list of the target sentences therefore consisted of 32 items.

2.2. Procedure

The experiment was run as a slide presentation (demo window) in Praat [2]. The target sentence was displayed as a series of pictures. Each subject, object and adverbial was schematically depicted. The participants were asked to memorize the pictures together with the word forms and to compose a sentence from the sequence of pictures and respond to the auditory presented context. Participants proceeded from slide to slide at their own pace with a mouse click.

2.3. Participants

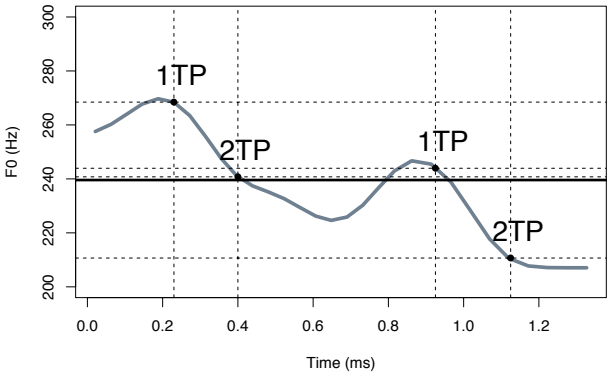
17 native speakers of Estonian (10 female, 7 male speakers) aged 22–40 years (m = 28.2 years) participated voluntarily in this study.

2.4. Analysis

The acoustic analysis of the recordings was carried out with Praat [2]. The F0 contour was manually segmented, relying on perception and visual observation of the F0 track.

In the nominal constituents of the sentence (grammatical subject and object), we determined the F0 maximum or – wherever a salient F0 maximum was missing – we annotated the beginning or the end of the excursion – *turning point* (TP, see Fig. 1).

Figure 1. Example of an annotation of the beginning (1TP) and end (2TP) of an F0 excursion. The black solid line represents the mean F0 of the speaker.



Four dependent variables were analysed: *peak height*, *slope, peak alignment* and *peak difference*. The *F0 peak* was either the F0 maximum or the F0 turning point. The *F0 slope* was the F0 change divided through the duration of the excursion. See the formula in (5) where F0 is given in semitones and *T* (ms) is the absolute time of the turning point.

(5) $\text{Slope (st/s)} = (F0_{1TP} - F0_{2TP}) / (T_{1TP} - T_{2TP})$

The *peak alignment* was the time of the peak or the first turning point in proportion to the duration of the first stressed vowel. See the formula in (6) (cf. the procedures in [22] and [16])

(6) $T_{prop} = 100(T_{F0}-T_{vOn})/(T_{vOff}-T_{vOn})$

where T_{F0} is the time of the F0 point, T_{vOn} the time of the vowel onset and T_{vOff} the time of the vowel offset. The *peak difference*, i.e. the difference

between the F0 peaks, was calculated by subtracting the F0 maximum or turning point (in semitones) of the second peak from the F0 maximum or turning point of the first peak.

For the corrective focus condition, we expect the peaks to be higher and later aligned and the F0 slope to be steeper compared to the new information focus condition. In addition, peaks are expected to be higher and later aligned and to have steeper slopes when the focus occurs on the object as opposed to the subject noun phrase. The sentence-initial object noun phrase is expected to attract nuclear prominence. This effect is expected to result in a greater peak difference for OVS word order than for SVO word order.

For the statistical evaluation, separate linear mixed models (lmer, [16]) were run with dependent variables peak height, slope, peak alignment and peak difference, and each with random factors speaker and item and with fixed factors focus type (corrective *vs.* new information focus), position (initial *vs.* final) and either grammar (subject *vs.* object) or word order (SVO *vs.* OVS).

3. RESULTS

Only sentences with high pitch accent (H*) were included; all sentences with low pitch accents (L*, 26 observations), incorrect word order or hesitations (77 observations) were omitted from the current analysis.

There were no significant interaction or main effects for focus type or grammar on peak height.

There were no significant effects of focus type or grammar on slope.

Figure 2. Peak alignment proportionally to the first stressed vowel separately for initial (light grey) and final (dark grey) sentence position, focus type (N = new information, C = corrective focus) and grammar (subject at the left side, object at the right side).

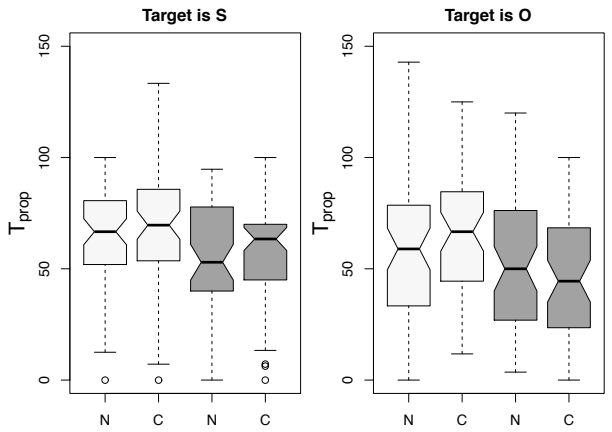
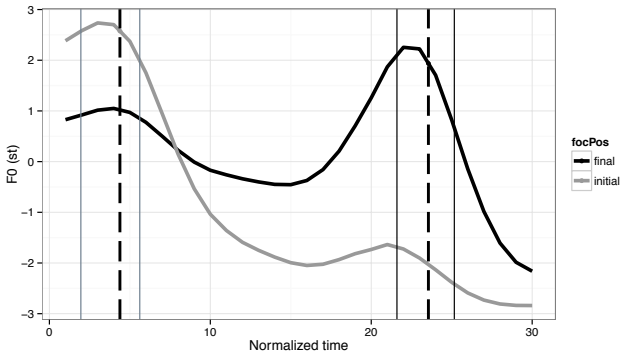
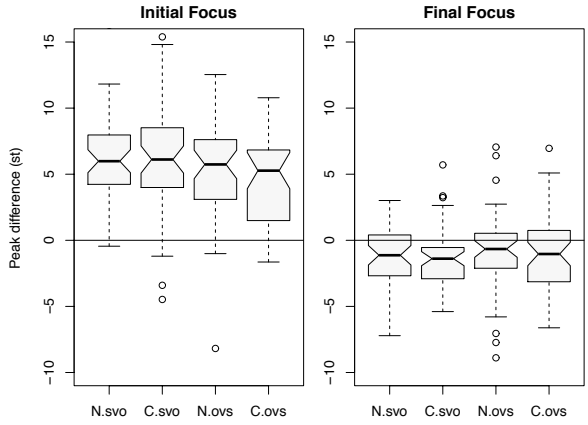


Figure 3. Averaged time-normalized F0 contours with sentence-initial (grey) and sentence-final (black) focus. Dashed vertical lines mark the location of the peak, solid vertical lines the boundaries of the stressed vowel of the word in focus.



As far as peak alignment was concerned, the proportional peak alignment varied between 39 and 78 ms but were unaffected by focus type or grammar. There was, however, a significant effect of sentence position on peak alignment. Commensurate with Fig. 2, sentence-final peaks (dark grey boxes) were aligned slightly earlier with the vowel than sentence-initial peaks (light grey boxes). As the time-normalized contours aggregated over all utterances and speakers show in Fig 3, the pitch peaks occurred on the word in focus that was located either at the beginning of the utterance (initial) or at the end of the utterance (final).

Figure 4. F0 peak differences (in semitone) separately for word order (SVO, OVS), sentence position of the focused word (left: initial, right: final) and focus type (N = new information focus, C = corrective focus)



Turning now to peak difference, this varied between -1.4 and 6 st (Fig. 4). The peak difference close to or below zero indicates that the peaks were either at the same height or that the first peak was slightly lower than the second peak. There were no

significant main effects of focus type or word order on peak difference and hence on declination. There was a significant interaction between word order and position of the focused word (χ^2[1]=7, p<0.01). A post hoc pairwise Tukey test showed a significant effect of position on peak difference (p<0.001): the peak difference varied between 3.6 and 7.3 st in phrases with sentence-initial focus, but between the -2.8 and 0.3 st in phrases with sentence-final focus. The significant effect of word order on peak difference (p<0.05) confirmed that the peak difference is smaller for OVS than for SVO word order – a result that runs counter our expectation.

4. DISCUSSION AND CONCLUSION

The study investigated intonational emphasis in terms of peak height, excursion size, peak alignment and peak difference in connection to two types of narrow foci. The main outcome of the study is that the type of focus affected none of the observed intonational variables. This result supports a model of Estonian by which narrow corrective focus does not necessarily cause additional intonational emphasis. Such a model is consistent with the model in [19] derived by a semantic analysis of focus.

In addition, even in sentence-initial object noun phrases speakers did not emphasize the word in focus. This finding, too, is consistent with the suggested model but can also be related to the possibility that the structural focus position might not need prosodic enhancement in Estonian. Another relevant outcome is the missing effect of grammatical function or word order on intonational variables, which implies that Estonian treats OVS word order similar to SVO in terms of narrow focus prosody. This is consistent with the model in [23] suggested for Finnish – another free word order language.

Studies for English also show inconsistent evidence [5, 11] for stronger intonational emphasis in corrective focus. It has been proposed that words in corrective focus form a logical contrast that may or may not be expressed in terms of a stronger accent ([3: 91]). The contradicting semantics of the context might induce the sense of greater prominence. It remains to be investigated whether the intuitively stated stronger prominence for correctiveness is a matter of acoustic reality or just the sensation of a particular pragmatic action, such as contradiction.

The peak difference did not reveal any difference between the two focus types. Thus the acoustic difference did not manifest in pre- or postfocal prosody as is proposed by [8]. However, the difference between correctiveness and new information may be perhaps realized on a rise preceding the fall or peak rather than as a difference between a nuclear and a pre- or post-focal peak. The study did not explore this possibility.

The peak was aligned earlier in phrase-final than in phrase-initial constituents. The peak alignment has been found to be affected by the following prosodic context [22, also recently for Estonian: 16]. One explanation is that the leftward shifting of a peak might come about because of the need to accommodate for a phrase-final low tone [22]. This effect is often called as *tonal crowding* (e.g. [13]). For example, [10] proposes that the earlier peak of a nuclear pitch accent reflects the tonal crowding in Spanish in which the phrase accent between the nuclear pitch accent and the boundary tone pushes the tonal target of the nuclear pitch accent to the left. We are not able to demonstrate on the basis of the current results whether the effect of position is caused by tonal crowding. It may be that a low boundary tone causes a form of tonal repulsion in which the pitch accent is aligned earlier in the phrase-final stressed vowel.

Another possible explanation for this finding may be phrase-final lengthening. Recall that, following [22, 16], the time of the peak was estimated in proportion to the length of the stressed vowel. A closer look into the vowel durations showed significantly longer vowels at the end of the phrase (av. 13.8 ms *vs.* 16.5 ms; χ^2[1]=45.2, p<0.001). The earlier peak alignment might therefore be the consequence of phrase-final lengthening that causes the absolute location of peak to be earlier (just as [22] concludes for English). This in turn implies that there might be some kind of tendency in Estonian to maintain a constant duration between the rise towards the peak or the plateau and the turning point. Further studies on phrase-level peak alignment are needed to provide a clearer picture of the Estonian intonational phonology.

The investigation presented in this paper revealed no acoustic difference between corrective and new information focus in Estonian. Whether such acoustic differences exist is also controversial in other languages. The interaction between acoustics, syntactic structures and pragmatics challenge further studies on prominence perception.

5. ACKNOWLEDGMENTS

The authors are grateful to Eva Liina Asu-Garcia and Pärtel Lippus for their valuable comments. The research was partly supported by Estonian Research Council grant IUT2-37.

6. REFERENCES

[1] Baumann, S., Grice, M. & Steindamm, S. (2006). Prosodic Marking of Focus Domains - Categorical or Gradient? *Proceedings from Speech Prosody*, 301–304

[2] Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer. Computer program.

[3] Bolinger, D. (1986). Intonation and its parts. Stanford, California: Stanford University Press.

[4] Breen, M., Fedorenko, E., Wagner, M. & Gibson, E. (2010). Acoustic correlates of information structure. *Language and cognitive processes*, 25: 7, 1044–1098.

[5] Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. Li (Ed.), *Subject and Topic* (pp. 67-78). London: Academic Press, Inc.

[6] Chen, Y., & Braun, B. (2006). Prosodic realization in information structure categories in standard Chinese. In R. Hoffmann, & H. Mixdorff (Eds.), *Speech Prosody 2006*. Dresden: TUD Press.

[7] Erelt et al., (1993). Eesti keele grammatika II. Süntaks. Tallinn: Eesti Teaduste Akadeemia, Keele ja Kirjanduse Instituut.

[8] Gussenhoven, C. (2007). Types of focus in English. In C. Lee, M. Gordon, D. Büring (Eds.) Topic and Focus. Cross-Linguistic Perspectives on Meaning and Intonation. *Studies in Linguistics and Philosophy*, 82, 83–100.

[9] Hanssen, J., Peters, J. & Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. Proceedings of Speech Prosody, 609-612. (Retrieved on 31. March 2014 from http://sprosig.isle.illinois.edu/sp2008/papers/id153.pdf).

[10] Hualde, J. I. (2002). Intonation in Spanish and the other Ibero-Romance Languages. Overview and status questions. In Wiltshire, C. R. & J. Camps (Eds.), *Romance Phonology and Variation* (pp. 101-115). Amsterdam: John Benjamins B.V.

[11] Katz, J. & Selkirk, E. (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. Language, 87(4), 771-816.

[12] Kohler, K., 1991. The perception of accents: F0 peak height versus F0 peak position. *AIPUK* 25, 219-241.

[13] Ladd, D.R. (2008). Intonational Phonology. Cambridge Studies in Linguistics, 119. Cambridge: Cambridge University Press.

[14] Ladd, D.R. & Morton, R. (1997). The Perception of Intonational Emphasis: Continuous or Categorical? *Journal of Phonetics*, 25, 313-342.

[15] Lindström, L. (2006). Infostruktuuri osast eesti keele sõnajärje muutumisel. (On the role of the information structure in the change of the Estonian word order). *Keel ja Kirjandus*, 11, 875–888. (Language and Literature, 11, 875–888).

[16] Plüschke, M. (2011). Peak alignment in falling accents in Estonian. *Proceedings of XVIIth International Conference of Phonetic Sciences*, Hong Kong 17-21 August.

[17] R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[18] Rietveld, A. C. M. & Gussenhoven, C. (1985). On the relation between pitch excursion size and prominence. Journal of Phonetics, 13, 299-308.

[19] Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75–116.

[20] Sahkai, H., Kalvik, M.-L. & Mihkla, M. (2013). Prosody of contrastive focus in Estonian. *Proceedings of the 14th Annual Conference of the International Speech Communication Association*, Lyon, France, August 25-29, 2013 (315-319). International Speech Communication Association.

[21] Salveste, N. (2013). Focus perception in Estonian: Is it governed by syntax or by prosody? In E. L. Asu & P. Lippus (Eds.), *Nordic Prosody – Proceedings of the XIth Conference*, Tartu 2012 (pp. 333-342). Peter Lang Verlag.

[22] Silverman, K., E., A. & Pierrehumbert, J. B. (1990). The timing of prenuclear high accents in English. In J. Kingston & M. E. Beckman (Eds.) Papers in laboratory Phonology I, Between grammar and Physics of Speech (pp. 72-106).

[23] Välimaa-Blum, R. M. (1993). Intonation: a distinctive parameter in grammatical constructions. *Phonetica*, 50, 124-137