Journal of Phonetics 65 (2017) 77–93

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

On the phrase-level function of f0 in Estonian

Nele Ots

University of Tartu, Estonia

ARTICLE INFO

Article history:
Received 27 November 2015
Received in revised form 9 June 2017
Accepted 23 June 2017
Available online 12 July 2017

Keywords:
Intonation
Narrow focus
Free word order
Estonian

ABSTRACT

In intonation languages, prominence-lending pitch movements are frequently aligned with sentence focus. In some free word order languages f0 may act as an acoustic correlate of focus together with focus-sensitive variation in word order. With respect to Estonian, a flexible word order language, this raises the question of how word order, which is sensitive to focus structure, interacts with intonation. This study explored how focus and word positioning affect f0 via a production experiment involving elicitation of laboratory speech. The results showed that independently of word order, the focal word was aligned with prominence-lending f0 movement, thus demonstrating that in Estonian narrow focus is related to f0, as is typical in intonation languages.

1. Introduction

This paper discusses how sentence intonation is used to signal sentence focus in Estonian – a Finno-Ugric language with relatively flexible word order rules. Word order variants in Estonian and other flexible word order languages have often been related to discourse relations such as sentence focus. It has long been recognised that free word order languages are also very often intonation languages and that intonation plays an important role in determining the meaning of a sentence (e.g. Daneš, 1960; Lenerz, 1977). However, in phonetic research there has been little interest in the interface between word order and intonation. The intriguing issue is how intonation interacts with sentence focus independently of word order.

Face and D'Imperio (2005) observed that in Italian and Spanish, sentence focus is conveyed by intonation as well as word order, perhaps even preferentially by intonation. Skopeteas, Féry, and Asatiani (2009) showed that in Georgian different word orders can be successfully produced with different intonation patterns. The results of their follow-up perception experiment confirmed that prosodically prominent words – not necessarily the words in the 'syntactic' focus position – are interpreted as the focus of the sentence. Välimaa-Blum (1993) found that in the case of Finnish, a Finno-Ugric language, different word orders can be uttered with different prosodic patterns depending on context and argued from this that the information-structural implications of word order can be overridden by intonation which, in turn, interacts directly with sentence focus. Vainio and Järvikivi (2007) provided corroborating evidence from a speech production study showing that speakers of Finnish increase the difference in prominence between two pitch accents to make the 'syntactic' focus position less prominent. These results demonstrate that the relationship between sentence focus and intonation is critical in Finnish and, probably, a number of other free word order languages.

1.1. Focus semantics and its interactions with pragmatics and syntax

This study relies on the focus theory put forward by Rooth (1992). The semantic function of focus is to indicate the presence of alternative interpretations of an utterance (Krifka, 2007; Rooth, 1992). To illustrate this, consider the sentence *MELANIE picked some blueberries* (henceforward capitals indicate the sentence focus). Focusing on the word *Melanie* informs the reader or listener that there were other people, who could also have picked some blueberries but did not. In accordance with its semantic function, focus can also induce additional pragmatic interpretations such as corrective, exhaustive or countersuppositional meanings of a constituent (see e.g. Gussenhoven, 2007; Krifka, 2007). These are referred to as focus types. In syntactic or phonological accounts of focus the distinction between narrow and broad

E-mail address: nele.ots@ut.ee

http://dx.doi.org/10.1016/j.wocn.2017.06.003
0095-4470/

focus is often important. *Narrow focus* refers to the single word (e.g. *Melanie* in the example above) and the *broad focus* to some larger unit, perhaps as the whole sentence (e.g. Melanie picked some blueberries) or the so-called 'verb phrase' (e.g. picked blueberries) in focus (see Gussenhoven, 2007; Ladd, 2008). The semantics of focus is consistent with the wide and narrow scope of focus (just as *bought some blueberries* is a potential alternative to *picked some blueberries*, so also is *picked some mushrooms*).

A number of other contextual factors have been shown to influence the linguistic output of a message. One is *information status* (also givenness), which can range from given to new based on previous mentions in a discourse: a previously mentioned concept is 'given' whereas a concept which has not been mentioned before is 'new' (Baumann, 2006; Baumann & Grice, 2006; Chafe, 1976; Prince, 1981). In English, but not necessarily in other languages, mentioned or inferable constituents are often placed before previously unmentioned constituents (e.g. Arnold, 2008; Clark & Haviland, 1977). Thus, information status very often interacts with referential means such as pronominal expressions and word order (e.g. Arnold, 2008). In theory, any constituent can be in focus regardless of its information status – new, partially new or given information – depending on the context (Baumann & Kügler, 2015). This implies that information status as signalled by word order may well be accompanied by sentence focus as indicated by intonation. There is, however, a statistical association between newness and focus: new information is frequently in focus (Krifka, 2007).

In addition to information status, sentences are known to reflect *theme–rheme* division in a manner which is analogous to topic–comment and topic–focus structures (see e.g. Daneš, 1966; Firbas, 1966; Gundel, 1985; Halliday, 1967; Lambrecht, 1994; Reinhart, 1981; Sgall, Hajicová, & Panevová, 1986). The theme can be thought of as the *part of an utterance which connects it to the previous context* (Steedman, 2000) or the part that speakers presume (or fail to presume) is common knowledge amongst the interlocutors (Steedman, 2014). The rheme is simply a comment on the theme (Halliday, 1967) or an update on the shared knowledge (Steedman, 2014). According to some researchers, the division into theme and rheme is insensitive to givenness or sentence focus and, in theory, a spoken sentence could contain a focus in addition to the logical organisation of message into matter and comment (Halliday, 1967; Steedman, 2000).

The most recent theoretical accounts claim that the information structure of a sentence can be represented in terms of at least two orthogonal dimensions (cf. (Baumann & Kügler, 2015; Calhoun, 2012; Krifka, 2007; Steedman, 2000)). This study adopts a theoretical framework that combines the phenomenon of sentence focus with the theme–rheme division. The theme–rheme division reflects the fact that, ideally, a sentence should comment on some matter (Halliday, 1967; Steedman, 2000). The concept of sentence focus accounts for the parts of a sentence that are relevant to the computation of contextual alternatives (Rooth, 1992). This study constitutes a phonetic investigation of the relationships between sentence focus, word order and intonation in Estonian.

1.2. Acoustic correlates of sentence focus

The main acoustic correlate of intonation is fundamental frequency (f0) – a rough estimate of the frequency of glottal pulses or a common denominator of multiple higher-level vibrations of a quasi-periodic speech wave. For a given air pressure (dB), the higher the f0, the louder the sound is perceived to be (Robinson & Dadson, 1956) and this is probably what underpins the common assumption that prominence is associated with a high f0 relative to the speaker's average f0 or f0 range. This means that a relatively high f0 is commonly associated with the syllables of a word and the words in a phrase that, to the non-expert listener, stand out – seem emphasised or prominent. In this section f0 is discussed in relation to prominence in terms of phrase or sentence accent (as represented by capital letters in the sentence *She brought WINE, not whisky*).

Research on the phonetics of prominence has produced conflicting results on whether f0 is a good correlate of prominence. Pioneering studies have assessed duration, f0 and intensity as possible correlates of prominence in American English and found that f0 was the most reliable correlate of the prominence of a stressed syllable (Fry, 1955, 1958; Lieberman, 1960). Further studies have demonstrated that the location of the highest pitch and largest pitch excursion in a phrase serve as the main correlates of focus in, for example, American English (Breen, Fedorenko, Wagner, & Gibson, 2010; Cooper, Eady, & Mueller, 1985), Dutch (Hanssen, Peters, & Gussenhoven, 2008; Swerts, Krahmer, & Avesani, 2002), Finnish (Suomi, Toivanen, & Ylitalo, 2003; Vainio & Järvikivi, 2007), German (Baumann, Grice, & Steindamm, 2006; Féry & Kügler, 2008) and Standard Chinese (Chen & Braun, 2006).

A few corpus studies have found evidence that contradicts the notion that f0 is a good correlate of prominence (Cole, Mo, & Baek, 2010; Kochanski, Grabe, Coleman, & Rosner, 2005; Turk & Sawusch, 1996). There are also some studies of laboratory speech comparing local measures of f0 (e.g. peak height, range of fall) in words elicited in a variety of contexts – such as broad vs. narrow focus – that found f0 to be a rather weak correlate of phrasal prominence. For instance, Baumann et al. (2006) showed that duration is a more general acoustic indicator of focal breadth than tonal properties, which tend to vary according to the speaker. Studies investigating the phonological notion of focus projection have also frequently attested that there is no tonal distinction between narrow focus on the sentence-final word and focus on the whole utterance (e.g. experimental demonstrations in Dutch and Greek by Gussenhoven (1983) and Gryllia (2009) respectively). It seems, therefore, that in at least some languages there may be no f0 rise on the word in focus relative to the broad focus.

A number of studies indicate that global rather than local f0 modulations across various contexts are important for sentence focus (Cooper et al., 1985; Swerts et al., 2002). For instance, phonetic studies of intonation have provided convincing evidence that the acoustic reflection of sentence focus is a post-focal drop in f0, accompanied by a narrow f0 range on non-focus words (Breen et al., 2010; Cooper et al., 1985; Xu

& Xu, 2005). Cooper et al. (1985) found that the increase in f0, and even the increase in the duration of words in focus, depended heavily on the position of the word in the phrase. Importantly, the f0 peak was not higher in narrow focus phrase-final words than broad focus phrase-final words. Similarly, the f0 peak at the beginning of the phrase containing the sentence-initial narrow focus did not differ from the corresponding peak in broad focus sentences. The magnitude of f0 rise in the phrase-medial focal word in narrow focus was still small (10 Hz) (Cooper et al., 1985, p. 2151). In addition duration, as an acoustic correlate of focus, depended on the phrase position of the word in focus: in the case of phrase-final words in focus it was 14.3%, whereas in other positions it was about 40% (Cooper et al., 1985, p. 2150). More importantly, however, (Cooper et al., 1985, p. 2151) observed a post-focus drop in f0. The considerable decrease in f0, accompanied by decreases in duration and intensity, is typically related to the perception of *deaccentuation* (see e.g. Cruttenden, 2006).

In the framework of autosegmental-metrical (AM) phonology (see e.g. Ladd, 2008), the prominence-lending f0 movement consisting of well-defined variance in f0 (e.g. f0 maximum as a peak, f0 minimum as the deepest point of a valley) that is associated with the stressed segment of a word is called the *pitch accent*. In very general terms, deaccentuation is the absence of a pitch accent. Thus, deaccentuation of non-focal words results in context-dependent differences in the number of pitch accents in a particular phrase. If a phrase has a narrow focus it has fewer pitch accents than if it has a broad focus. For instance, Swerts et al. (2002) observed that in Dutch an utterance with phrase-initial narrow focus contains fewer pitch accents than an utterance with phrase-final narrow focus. This phenomenon could be referred to as the context-dependency or focus-driven distribution of pitch accents. This paper starts from the idea that distribution of pitch accents is a much more reliable indicator of focus than local f0 rise or any other local expansion of f0 excursion. The important implication for the analysis of f0 is that local features, such as peak height and range of excursion, need to be analysed alongside the global features of an intonation phrase.

One of the well-known global features of phrasal pitch is *declination* – a gradual downwards drift in f0 within an intonation phrase (see e.g. Cohen, Collier, & t'Hart, 1982; Gussenhoven, Repp, Rietveld, Rump, & Terken, 1997; Pierrehumbert, 1979; Yuan & Liberman, 2014). Declination is usually modelled as linear regression lines fitted to local f0 peaks or valleys, in which case they are called *topline* and *baseline* (Cohen et al., 1982; Yuan & Liberman, 2014), or to all f0 values making up the contour, in which case they are called overall regression lines or midlines (e.g. Liberman & Pierrehumbert, 1984; Swerts, Strangert, & Heldner, 1996). The topline slope can be exploited for measuring the relative *scaling* of a pitch peak: a peak might be higher or lower in pitch than the preceding or following peak, which would result in positive or negative topline slopes respectively. A related fact is that the declination of the phrase-final peak has been shown to interact with perception of relative prominence. A lower pitch peak towards the end of a phrase, is perceived as having the same prominence as the preceding higher peak (Gussenhoven et al., 1997; Pierrehumbert, 1979). Thus, if the sentence-final word in focus is to be signalled by high f0 peak, the height of the peak does not necessarily need to exceed the absolute height of the previous peaks.

In summary, the phonetic study of theme–rheme division and sentence focus in Estonian should include local acoustic measures (e.g. range of f0 on a particular word, duration of a particular word) as well as global acoustic intonation characteristics (f0 decline).

1.2.1. Linguistic functions of f0 in Estonian

A well-investigated phenomenon in relation to the function of f0 in Estonian is the three categories of relationship between stressed and unstressed syllables – three-way *quantity* distinction. Estonian words can contain either first quantity (Q1, short), second quantity (Q2, long) or third quantity (Q3, overlong) that are distinguished by the duration ratio of the stressed and the unstressed syllables (for further details and examples see Lehiste, 1960, 1997; Lippus, Asu, Teras, & Tuisk, 2013). As the durational difference between Q2 and Q3 may be acoustically quite subtle, the quantity contrast can also be distinguished prosodically – by the timing of the pitch peak (Eek, 1974; Lehiste, 1960; Lippus et al., 2013; Mihkla & Kalvik, 2011). In a Q2 word pitch peak is aligned with the second half of the stressed syllable, whereas in a Q3 word it is aligned with the first half of the stressed syllable (see Fig. 1). This tonal feature has been shown to be phonological in nature (Plüschke, 2013) and critical to native quantity recognition (Eek, 1980, 1983; Lehiste, 1997; Lippus, Pajusalu, & Allik, 2009, 2011; Salveste, 2010).

There have been few studies of the sentence-level functions of f0 in Estonian. Some studies have described Estonian pitch accents and phrase boundary tones using an autosegmental-metrical framework (Asu, 2004, 2005; Asu & Nolan, 1999, 2007), strongly implying that Estonian is an intonation language much like English or German. According to Asu (2005), the distribution of pitch accent types in Estonian intonational phrases can be different in nuclear and pre-nuclear positions. Pre-nuclear positions can contain a high falling pitch accent (H*+L) or a downstepped high pitch accent (H+!H*). Nuclear positions can be associated with high or low accents (H*, L*+H, L*) or a high falling accent with an early fall and a low target in the stressed vowel (H+L*). The early fall (H+L*) has the interesting property that it necessarily recurs if it is introduced: if there is an early fall at an accentable position, then all the subsequent accentable positions must contain

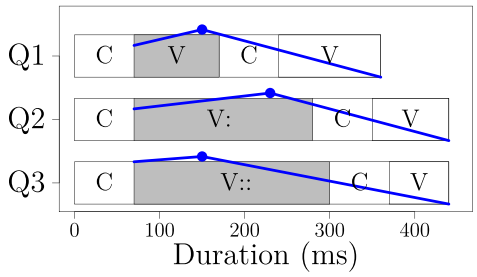

Fig. 1. Quantity schema showing tonal features (blue lines) and the approximate segment durations. The grey bar demonstrates the location of the stressed vowel; C and V mark the boundaries of consonants and vowels respectively. The blue circles mark the pitch peak and the blue lines the overall pitch movement relative to stressed vowel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

early falls (Asu & Nolan, 2007). Asu (2004, 2005)found no evidence for the existence of intermediate phrase boundaries (ip). Therefore, only intonational boundaries (IP) have been modelled for Estonian intonational phrases. High or high falling accents are followed by a falling boundary tone or a sustained boundary tone (%), whereas low accents (L*+H, L*) are followed by a high boundary tone (H%).

It is less clear what the functions of these intonational categories are. High or rising boundary tones have been shown to have a pragmatic function: in conversations they are used as turn-holding signals (Asu, 2006). Sahkai, Kalvik, and Mihkla (2013b) found that a word in focus usually carries some kind of pitch accent, most frequently a high pitch accent (H*). Some recent studies have investigated whether sentence focus is associated with pitch accent, f0 rise, longer duration and greater intensity (Mihkla, Sahkai, & Kalvik, 2015; Sahkai, Kalvik, & Mihkla, 2013a, 2013b). Sahkai et al. (2013a) reported, for instance, that focus has more impact on word duration than on f0.

The latest study by Mihkla et al. (2015) investigated production of sentences with a broad or narrow focus on the last word of a phrase. They found that f0 was the weakest correlate of sentence focus amongst the set of acoustic variables they investigated (f0 range; alignment of the turning point; mean intensity; spectral emphasis). On the basis of this result Mihkla et al. (2015) concluded that in Estonian f0 does not signal focus; this is consistent with the view that intonation is redundant or even unnecessary in a language which allows information structure to be encoded in word order. The result of Mihkla et al. (2015) is, however, unsurprising, because studies of other languages have already indicated that the broad and narrow focus on the last word of a phrase are not consistently distinguished by tonal features (e.g.Cooper et al., 1985; Gryllia, 2009; Gussenhoven, 1983).

1.2.2. Estonian word order and its interactions with sentence focus

From the perspective of phrase structure grammar, Estonian has free word order (Ehala, 2006; Erelt et al., 1993; Lindström, 2006; Remmel, 1963; Vilkuna, 1998), but this is relative since word order is sometimes highly context-dependent (e.g. Ehala, 2006; Erelt et al., 1993; Lindström, 2006; Remmel, 1963; Vilkuna, 1998). Although different word orders are quite frequent in the written language corpora (Tael, 1988) and spontaneous speech (Lindström, 2002, 2004, 2006), their semantic properties are not well understood. According to Erelt et al. (1993) the word order variants SVO (subject-verb-object) and OVS arise due to the theme–rheme structure (see Examples 1 and 2).

(1) Peeter luges “Sõrmuste isandat”
Peter read.SG3.PST. ring.PL.GEN lord.SG.PRT
‘Peter read “Lord of the Rings”’
(2) “Sõrmuste isandat” luges Peeter
ring.PL.GEN lord.SG.PRT read.SG3.PST. Peter
‘Peter read “Lord of the Rings”’

The difference between examples (1) and (2) is that in (1) the rheme is the object of the action (“*Lord of the Rings*”) because the sentence is most compatible with the question *What did Peter read?* whereas the rheme in (2) is the agent of the action (*Peeter*) because the sentence is more compatible with the question *Who read the Lord of the Rings?* This analysis is supported by the corpus studies that have reported that sentence-final subject noun phrases (NPs) are quite frequent in the corpus of spontaneous speech and mostly contain an as yet unmentioned referent (Lindström, 2002, 2004).

Syntactic research has not provided any evidence that there is an obligatory focus position in Estonian, as there is in Hungarian, another Finno-Ugric language (see in Kiss, 1995; Siptár & Terköczy, 2000; Szendröi, 2003). The results from the experiment described by Sahkai et al. (2013b) could be taken as empirical support for the absence of obligatory focus position. Sahkai et al. (2013b) asked speakers to answer a number of questions about the pictures that depicted scenes involving the agent and the object of an activity. Although speakers were free to choose the word order for their full-sentence responses, they used pitch accent, rather than word order, to indicate sentence focus. The fact that speakers did not vary the positioning of words to signal focus indicates that word order is insensitive to focus and is instead used to indicate thematic relations (as illustrated in Examples 1 and 2 above). In addition, Erelt et al. (1993) cite intonation as a means of indicating sentence focus and so in their tentative model of Estonian information structure the theme–rheme structure is not related to sentence focus.

There is no easy way to keep thematic relations and sentence focus separate from each other, partly because in the literature the presence or absence of a particular structure (sentence focus or theme–rheme division) in a given utterance is typically diagnosed using the same method — by accommodating a *wh*-question to an existing sentence or utterance. To illustrate, a sentence *MELANIE picked some blueberries* is usually analysed as being the most compatible with a question *Who picked some blueberries?* According to some linguistic analyses, the correspondence between *wh*-phrase and the pitch-accented word *Melanie* indicates that the proper name Melanie is the sentence focus (Halliday, 1967; Lambrecht, 1994; Rooth, 1992; Sgall et al., 1986) but by some other analyses it can be also the rheme of a sentence (like in Sgall et al., 1986; Steedman, 2000, 2014). One way to distinguish the effects of theme–rheme structure and sentence focus is to embed the contrast into theme–rheme structure (see Calhoun, 2012). The interaction between theme–rheme structure and sentence focus in Estonian is illustrated in example (3) (the sentence-initial constituent is always the (th) eme and the rest of the sentence the (rh)eme; the sentence focus is indicated with capitals).

(3) **Q1:** Kas Liina luges “Sõrmuste isandat”?
(‘Did Liina read “Lord of the Rings”?’)
A1: Ei, [PEETER]$_{th}$ [luges “Sõrmuste isandat”]$_{rh}$.
A2: Ei, [“Sõrmuste isandat”]$_{th}$ [luges PEETER]$_{rh}$.
(‘No, PETER read “Lord of the Rings”’)
Q2: Kas Peeter luges “Tõde ja õigust”?
(‘Did Peter read “Truth and Justice”?’)
A3: Ei, [Peeter]$_{th}$ [luges “SÕRMUSTE ISANDAT”]$_{rh}$.
A4: Ei, [“SÕRMUSTE ISANDAT”]$_{th}$ [luges Peeter]$_{rh}$.
(‘No, Peter read “LORD of the RINGS”’)

In example (3), the different versions of answers (As) to the questions Q1 and Q2 vary in word order: A1 and A3 follow SVO word order whereas A2 and A4 follow OVS word order. The combination of Q1 and answers A1 and A2 implies a

contrast between agents Liina and Peeter; the combination of Q2 and answers A3 and A4 implies a contrast between the two novels. Both word orders should be compatible with both questions provided the appropriate constituent is accented (shown with capital letters; inferred from Erelt et al., 1993). Answers in A1 and A4 may indicate that the concepts in focus relate to the previous context and are a shared knowledge amongst the interlocutors, whereas the answers A2 and A3 indicate that the concepts in focus have not been introduced yet and do not belong to the common background.

Another way to separate theme–rheme structure and sentence focus would be to say that utterances that are underspecified for narrow focus or for an underlying *wh*-question (referred to as all-rheme sentences or sentences with unmarked themes in Steedman (2000, 2014)) naturally represent the theme–rheme structure (inferred from Halliday, 1967). In this case, the pure theme–rheme structure of the sentences in (1) and (2) would emerge in the context of a question such as *What happened?*

1.3. Objective of the experiment

Sahkai et al. (2013b) found that a word in focus frequently carries a high pitch accent (H*) and so this study proceeds from the idea that f0 is the main correlate of phrase-level prominence in Estonian. Mihkla et al. (2015) found that f0 is not a reliable acoustic correlate of phrasal prominence in Estonian because narrow focus and broad focus of sentence-final words did not differ in f0 range.

This study extended earlier research (Sahkai et al., 2013b; Mihkla et al., 2015) by including the narrow focus of a phrase-initial word and the narrow focus of a phrase-final word. My hypothesis was that in the narrow-focus condition, the constituents in focus would have a pitch accent whereas the non-focal constituents would be deaccented, in the sense that the f0 excursions aligned with them would be significantly lower and narrower in pitch range than excursions in the pitch accent condition.

The experiment was also designed to explore the interaction between sentence intonation and word order. Sahkai et al. (2013b) were not able to investigate the effects of word order on sentence intonation because the speakers in their experiment did not vary the position of the word in focus. Here I present data collected in a more controlled context in which participants were not able to choose the word order for a transitive sentence. Despite of such a controlled experimental context, the results enabled me to draw conclusions about the possible limits of the intonational structure of a language, especially in terms of prosodic sentence focus. The hypothesis about the interaction between sentence intonation and word order was that the focus of a sentence is aligned with prominence-lending f0 movement independent of word order (SVO vs. OVS); this was based on an assumption that positioning of a word interacts with the theme–rheme division, as was demonstrated in Section 1.2.2. As was also demonstrated in Section 1.2.2, both word orders are pragmatically plausible in various broad and narrow focus contexts.

The sections below present a phonetic investigation of the interactions between intonation, word order and sentence focus in Estonian. The experiment was designed to answer the two following questions.

1. Does narrow focus affect the relative prominence of the pitch peaks within an intonation phrase?
2. Does the interaction between word order and sentence focus influence the relative prominence of the pitch peaks?

2. Experiment

2.1. Method

The above questions were investigated in an experiment involving a speech elicitation task in which participants were asked to respond to a question played to them through headphones while they were viewing a series of pictures. Participants' speech was recorded and subjected to acoustic analysis to determine f0 and word duration.

2.1.1. Materials

Four *target sentences* were constructed. Target sentences consisted of three disyllabic words consisting mainly of sonorous sounds, see Table 1 below. The subjects were female names.

The order of the sentence constituents in the target sentence was either SVO (subject noun phrase (NP), verb then object NP) or OVS (object NP, verb then subject NP). Thus, all subject and object NPs occurred at the beginning or end of a sentence. The noun phrases at the beginnings and ends of the sentence were the measurement targets.

The noun phrases included an unbalanced number of second and third quantity because the experiment was based on the assumption that the tonal feature of word quantity would not interfere with sentence intonation. This assumption was derived from studies of intonational categories that suggest that Estonian is an intonation language (Asu, 2004, 2005; Asu & Nolan, 2007). From these studies I deduced that at phrase level, appropriate tonal encoding of sentence focus – if signalled by pitch accent – might be more important than tonal indication of quantity. This assumption is also supported by earlier research on the phonetics of the low pitch accent containing an early fall (H*L) (Asu & Nolan, 2007) and on interaction between tonal features and rising boundary tones (Asu & Nolan, 1999). Thus, the two target words consisted of third quantity (Q3), *vaala* and *raami*, and the others consisted of second quantity (Q2).

The target sentences in Table 1 were elicited as responses to the three types of *wh*-questions corresponding to the meaning of the target sentence. The first question was a question about the whole event and was used to elicit a so-called broad focus sentence. The second and third questions were wh-questions that were signalled by intonation and were designed to elicit narrow focus on the first word of the target sentence (focus-initial sentence) or on the last word of the target sentence (focus-final sentence). This way of presenting the *wh*-question was chosen because the participants in the pilot experiment found it very unnatural to give a full-sentence response to questions that explicitly included a explicit *wh*-phrase (see Table 2).

The design consisted of two word orders (SVO, OVS) crossed with three focus structures (broad, focus-initial,

Table 1
Lexical material used in elicited speech production experiment.

	Subject	Verb	Object	Translation
1.	Leena [l e: .n ɑ	maalis m ɑ: .l i s	vaala v ɑ:: .l ɑ]	'Leena drew a whale.'
2.	Liina [l i: .n ɑ	liimis l i: .m i s	raami r ɑ:: .m i]	'Liina repaired a frame.'
3.	Meeli [m e: .l i	hüüdis h y: .t i s	Loonat l o: .n ɑ t]	'Meeli called for Loona.'
4.	Miili [m i: .l i	kuulis k u: .l i s	Eevat e: .v ɑ t]	'Miili heard Eva.'

Table 2
Examples of questions inducing different focus structures in the sentence *Leena maalis vaala* 'Leena drew a whale.'

	Focus structure	Question
(a)	Broad focus	Mis juhtus? 'What's up?'
(b)	Focus-initial	Keegi ju maalis vaala? 'Who drew a whale?' lit. 'Somebody drew a whale?'
(c)	Focus-final	Leena ju maalis midagi? 'What did Leena drew?' lit. 'Leena drew something?'

focus-final), yielding 6 conditions for each of the four sentences in Table 1, hence there were 24 target sentences. The target sentences were supplemented by 24 *filler sentences* consisting of an object noun phrase (O), a verb (V) and an adverb (A) (e.g. *Laadal müüdi lilli*, 'The flowers were sold on the market') that also varied with respect to constituent order (AVO vs. OVA). Both types of sentences were presented together with three types of context in a random order that varied across participants. The distance between the instances of a given sentence was not controlled.

2.1.2. Procedure

The experiment was run as a slide presentation (demo window) in Praat (Boersma & Weenink, 2013). Participants advanced from one slide to the next by clicking a mouse and the experiment was self-paced. The context question was presented orally when the slide changed and was also displayed in the upper part of the screen. To elicit non-read speech whilst controlling word order, the sentences were displayed as a series of schematic pictures of the words (see Fig. 2). In a practice session carried out before the experiment, participants were asked to memorise the pictures and matching word forms. The task was to respond to the oral context question with a sentence in which the word order matched the order of the pictures on the screen.

2.2. Participants

The participants were 10 female and 7 male speakers ($N = 17$) aged between 22 and 40 years (mean age = 28.2 years). They all reported having normal or corrected-to-normal vision and no hearing difficulties. Their participation was voluntary.

2.3. Analysis

Acoustic analysis of the recordings was carried out with Praat (Boersma & Weenink, 2013). To enable analysis of

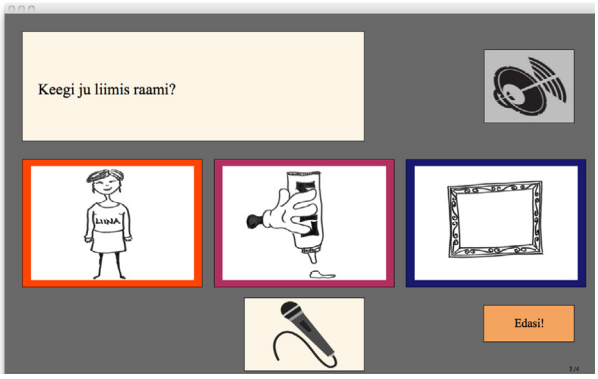

Fig. 2. Example slide. One of the three questions (e.g. What did Liina repair?) and a button with an image of a loudspeaker was presented in the upper part of the screen. The participant would hear the question and could read and replay it. The schematic pictures corresponding to the words the participant had to utter were displayed in the middle part of the screen, in the order in which the words were to be uttered. A button showing an image of a microphone was displayed in the bottom part of the screen. When the participant was ready, s/he had to click on the microphone image, utter the sentence and then advance to the next slide by clicking on the button marked Edasi! ('Continue!').

phrasal prominence the f0 track was tagged manually, on the basis of auditory and visual observation of f0 excursions. The high (H) and low (L) points of the excursion were marked. The high point of the excursion was the f0 maximum. If there was no clear f0 peak, then the approximate location of the *elbow* – the start of an abrupt negative pitch change, often called the *plateau offset* (for analogous analysis see e.g. D'Imperio, 2000; Knight, 2003) – was determined by reference to the vertical line under the cursor provided in the Praat analysis window. The low point of the excursion was taken as the lowest f0 towards the end of the word because annotation of the high and low points of the pitch excursion was restricted to locations within the target word, even when the pitch movement continued into the verb. In the case of very small pitch excursions (an almost flat contour), the two points were marked somewhere in the middle of the stressed and unstressed vowel. Thus, the two points defining the ends of the f0 excursion were marked in every target word, meaning that four points were marked in each phrase. A few examples of the annotation procedure are shown in Fig. 3.

The f0 values collected at the pitch maxima and minima or elbows were converted into semitones according to Formula 1

$$12 \cdot \log\left(\frac{f0(Hz)}{ref(Hz)}\right) = semitone(st) \quad (1)$$

where the reference value was the speaker's average f0, based on all the f0 samples from his/her utterances. Semitone

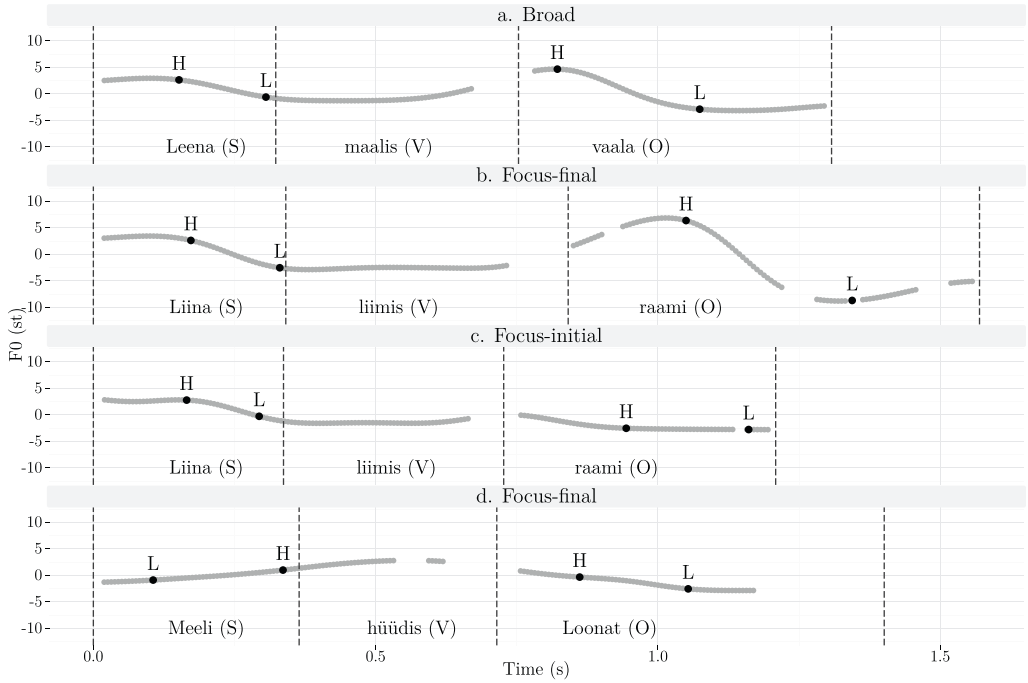

Fig. 3. Annotation of f0 excursions in the sentence-initial and sentence-final words. The long-dashed vertical lines indicate word boundaries. The black circles indicate the high (H) and low (L) points of the f0 excursion.

conversion was used because it partially filtered out gender-specific variation in f0 range.

Fig. 3 shows four examples of f0-trajectories from four different female speakers, in semitones. The pitch contour in the upper panel (a) comes from the first target sentence (*Leena maalis vaala*, see Table 1 above), the pitch contours in the second and the third panels (b and c), come from the second target sentence (*Liina liimis raami*, see Table 1 above) and the pitch contour in the bottom panel (d) comes from the third target sentence (*Meeli hüüdis Loonat*, see Table 1 above). The contours represent responses to different questions: in the upper panel *What happened?* (broad), in the second panel *What did Liina repair* (focus-final), in the third panel *Who repaired the frame?* (focus-initial) and in the bottom panel *Who did Meeli call?* (focus-final, see Table 2 for contexts). Plateaux-like realisation of high pitch accent can be observed at the beginning of contours (b)–(d). This phonetic realisation of high falling pitch accent (H*) was very frequent in the recordings and hence I developed the technique described above for identifying and annotating the elbows.

To ensure that only reliable data were analysed recordings exhibiting technical errors, utterances using the wrong word order or extremely creaky voice (9%) were omitted from the evaluation; this left 372 out of 408 utterances for further examination. High falling pitch accents were the most frequent accents, as can be seen from the examples in Fig. 3, although there was some variation in intonational patterns. For example, there were seven instances of the *hat pattern* (H* H*+L L%, see the contour in panel (d)), three instances of clear list intonation (f0 rise throughout the phrase) and two instances of recurring low accents (H+L*). Because these 12 utterances would have required a slightly different analytical approach they were omitted from the analysis, leaving a total of 360 utterances (88%) for evaluation. The exclusions were consistent with the goal, which was to examine the effect of focus structure (broad focus; focus-initial; focus-final) on the pitch range of the high pitch accent (H*/H*+L) in Estonian sentences with a SVO or OVS word order.

2.4. Results

2.4.1. Fundamental frequency

The time-normalised f0 contours in Fig. 4 are intended to give a rough overview of the pitch patterns and ranges (in semitones) of the sentences uttered for the different focus structures (broad; focus-initial; focus-final). They show that when the focus is broad or final there is a pitch peak on the last word of the sentence, whereas when the focus is initial there is a pitch peak on the first word of the sentence. In addition, a pitch peak for initial focus is followed by a decline in f0 across the rest of the phrase and the f0 excursion on the last word appears to be very low and narrow in range. Clearly, narrow focus phrases contain fewer prominence-lending f0 movements than broad focus phrases. In narrow focus contexts (focus-initial; focus-final) these effects do not seem to depend on word order (SVO vs. OVS); however word order does appear to affect the pitch pattern in broad focus sentences: the last pitch peak is lower in OVS than in SVO. Because time-normalised pitch contours do not contain any time-stamps the representations of peak alignments and other phenomena in Fig. 4 are only impressionistic. The effects observed in Fig. 4 can be quantified in terms of (a) the semitone range of excursion; (b) the f0 slope of the f0 excursion; (c) the timing of the f0 peak relative to the duration of the stressed syllable; (d) the duration of the word and (e) the baseline and topline slopes of linear fits to the high and low points of manually annotated f0 excursions.

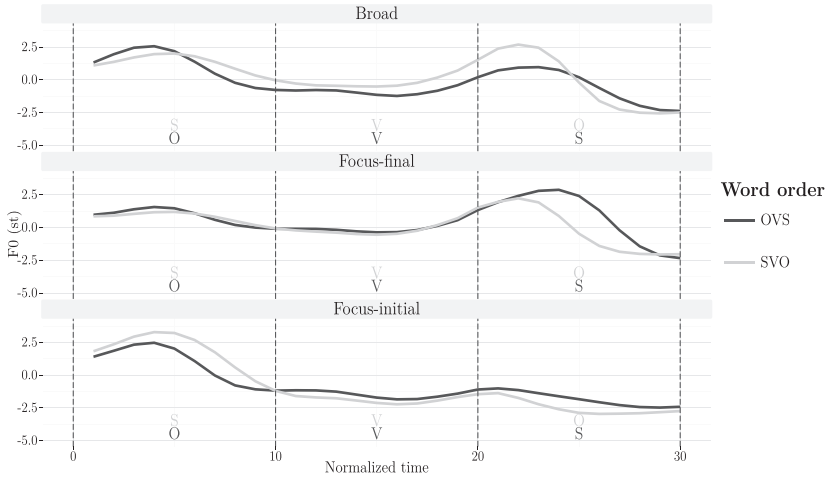

Fig. 4. Time-normalised f0 contours aggregated over 17 speakers and 4 target sentences as a function of focus position and word order. Ten f0 values per word were extracted from the three-word-utterances. The long-dashed vertical lines indicate word boundaries.

Range of excursion was defined as the difference between the high and low points of the f0 excursion in semitones. Slope was calculated using the formula in (2) below (adapted from Rathcke & Harrington, 2006).

$$\frac{f0_{1st} - f0_{2nd}}{T_{1st} - T_{2nd}} = slope(st/s) \quad (2)$$

where the f0 range (st) is divided by the duration (s) of the pitch excursion. Slope (st/s) was used to normalise the large f0 ranges possible over long time spans. For instance, an f0 fall of six semitones over 0.3 s has a slope of 20 st/s; if a six-semitone fall takes only 0.12 s then the slope is 50 st/s. Although the range is the same (six semitones) the slope is greater in the latter example. It is assumed that a faster rate of change (50 st/s rather than 20 st/s) has greater psycho-acoustic salience.

Three explanatory factors were defined for the purposes of the evaluation: (a) word order (two levels: SVO; OVS), (b) focus structure (three levels: broad focus; initial focus; final focus) and (c) the position of the word in the intonation phrase ((phrase-) initial; (phrase-) final). The phrase position factor captures whether a word occurred at the beginning or the end of the sentence: in the case of SVO sentences, the sentence-initial word was a subject NP but in the case of OVS sentence it was an object NP. Generalised linear mixed models with random slopes for subject (as a method available in the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R, R Core Team, 2017) were used to evaluate the effects of word order, focus structure and phrase position on the various dependent variables. *P*-values were obtained by calculating likelihood ratios for comparisons between the full model with a given interaction and the model without the interaction.

The effects of word order, focus structure and phrase position on pitch range are shown in Fig. 5. Fig. 5 shows that the pitch range is quite compressed in pre-focal positions (the phrase-initial word of the focus-final sentence; shown in light grey) and post-focal positions (the phrase-final word of the focus-initial sentence; shown in white) independently of word order. Pre-focal pitch range still is slightly greater than post-focal pitch range. In addition, the pitch fall on the first word of the sentence is greater in initial focus than in broad focus sentences (compare the dark grey and white boxplots for the initial position) irrespective of word order (SVO vs. OVS). There is, however, no difference between sentence-final focus and broad focus with respect to the last word of an utterance (compare the dark grey and light grey boxplots for the final position). As one of the reviewers noted, there is a lot of variance in pitch range in final positions. I find it very plausible that this is due to a number of the speakers producing rather narrow-ranged f0 excursions in that position. Finally, the effect of word order is observable only in the broad focus condition (dark grey): the phrase-final word of an OVS utterance has a slightly smaller pitch range than the phrase-final word of a SVO utterance.

The generalised linear mixed effects analysis with range as the dependent variable showed a three-way interaction between the three fixed effects (word order, phrase position and focus structure) ($\chi^2[2] = 17$, $p < .0001$). *Post hoc* Tukey tests showed that phrase position (phrase-final vs. phrase-initial) affected range in broad focus SVO sentences ($p < .01$), focus-final SVO sentences ($p < .0001$), focus-initial SVO sentences ($p < .0001$), focus-final OVS sentences ($p < .0001$) and focus-initial OVS sentences ($p < .0001$) but not broad focus OVS sentences. There was a difference between broad focus and initial focus SVO sentences with respect to range in the phrase-initial and phrase-final words (both *p*s < .0001). Notably, there was no difference between broad focus and final focus SVO sentences with respect to range in the phrase-initial and phrase-final words. There was a difference between broad and initial focus OVS sentences with respect to range in the phrase-final word ($p < .0001$), but not the phrase-initial word. Final focus OVS sentences differed from broad focus OVS sentences with respect to the range of the phrase-final and phrase-initial words ($p < .0001$ and $p < .05$ respectively). Word order (SVO vs. OVS) only affected range on the last word of the broad focus OVS sentence ($p < .05$).

As argued above, pitch range relative to the duration of the fall (slope) might be a better estimator of tonal prominence, and so the influence of focus structure and word order on slope is presented in Fig. 6.

The slope is greatest for narrow focus conditions (initial; final) (mean slope varies between 26 and 27 st/s, sd between 17 and 18 st/s) independently of sentence position. The pitch

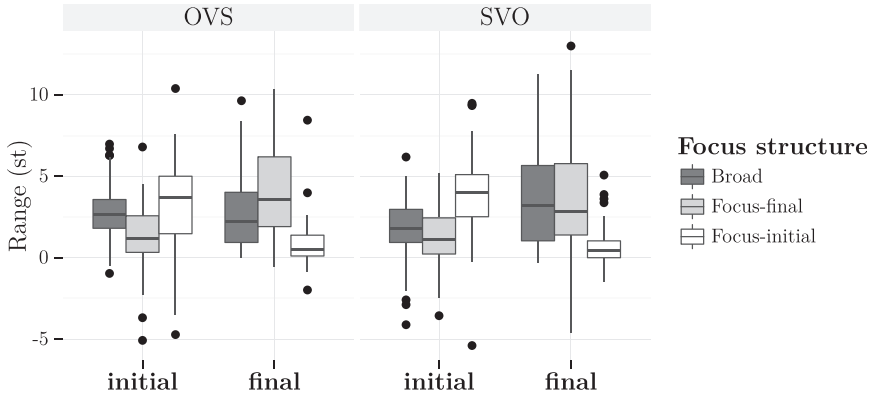

Fig. 5. f0 range (in semitones, see Formula 2) as a function of word order (SVO vs. OVS), focus structure (broad, initial vs. final focus) and phrase position (initial, final).

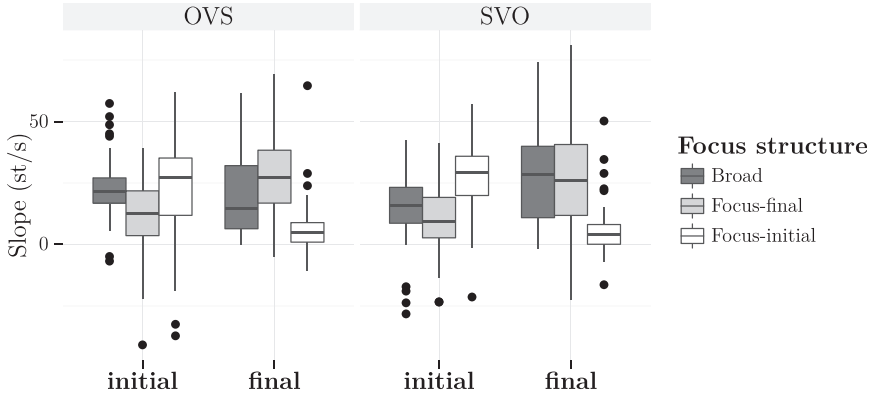

Fig. 6. f0 slope as a function of word order (SVO; OVS), focus position (broad; initial; final) and phrase position (initial; final).

excursion on non-focal words is influenced by phrase position. It is small or almost non-existent at the end of the sentence (mean = 6 st/s; sd = 10 st/s), but slightly greater at the beginning of the sentence (mean = 10 st/s; sd = 16 st/s). The slope in the sentence-initial position is greater in focus-initial sentences (mean = 28 st/s, sd = 20 st/s) than in broad focus sentences (mean = 20 st/s, sd = 17 st/s). In contrast, the slope in sentence-final words does not differ between broad and narrow focus sentences. The slopes in broad focus utterances appear to vary with word order: the slope on sentence-final objects (mean 27 = st/s; sd = 17 st/s) is greater than that on sentence-final subjects (mean 20 st/s; sd 21 st/s) in broad focus sentences.

The results from the generalised mixed effects analysis of slope are very similar to those for range. There was a three-way interaction between the fixed effects (word order, focus structure and phrase position) ($\chi^2[2] = 16$, $p < .0001$). *Post hoc* Tukey tests showed that phrase position affected slope in broad focus SVO sentences ($p < .001$), focus-final SVO sentences ($p < .0001$), focus-initial SVO sentences ($p < .0001$), focus-final OVS sentences ($p < .0001$) but not OVS broad focus sentences. There were differences between the slopes associated with broad and initial focus in the phrase-initial and phrase-final words of SVO sentences (both $ps < .0001$) but in the case of OVS sentences the slopes associated with broad and final focus only differed in the phrase-final word ($p < .0001$) and not in the phrase-initial word. Broad and final foci were associated with different slopes in the first word of the OVS sentence ($p < .05$) but not the last word of the OVS sentence or with any position in the SVO sentence. There was no effect of word order on slope.

The materials for the experiment were constructed so that the sentences would feature very sonorous sounds (long vowels). In Estonian, the difference between Q2 and Q3 (see Section 1.2.1) often encodes grammatical function of a word (object NP vs. subject NP), so inclusion of Q2 and Q3 words in object NPs was unavoidable and hence the design was unbalanced with respect to quantity. There were 546 instances of Q2 and 174 instances of Q3. It was assumed, however, that the tonal characteristics of quantity do not interact with focus structure. In the following, this assumption was also tested using linear mixed effects analysis. Slope as a function of quantity (Q2 and Q3) is shown in relation to phrase position (phrase-initial; phrase-final) and focus structure (broad; focus-final; focus-initial) in Fig. 7.

The slopes in Fig. 7 follow the same trends as the slopes in Fig. 6. They show effects of phrase position and focus structure, indicating that quantity does not influence the slope. A generalised linear mixed model with slope as the dependent variable and fixed effects of quantity (Q2; Q3), phrase position (phrase-initial; phrase-final) and focus structure (broad; initial; final) was defined. There was no three-way interaction between the fixed effects, so separate models were used to test for two-way interactions between quantity and focus structure and between quantity and phrase position. There was an interaction between quantity and focus structure ($\chi^2[2] = 12$,

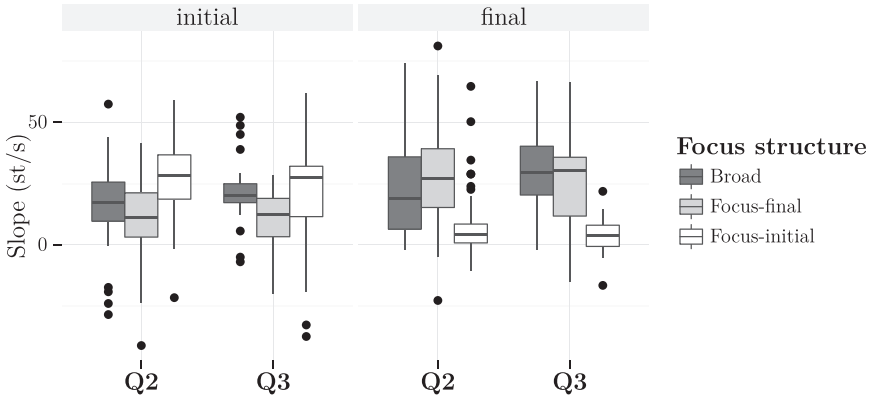

Fig. 7. f0 slope (semitones in seconds) as a function of quantity (Q2 vs. Q3), focus structure (broad; initial; final) and phrase position (initial; final).

$p < .05$). *Post hoc* Tukey tests showed that the slopes associated with broad and initial focus differed in Q3 words ($p < .0001$) but not in Q2 words.

I discussed the timing of the f0 peak in relation to duration of stressed syllable in Section 1.2.1 and Fig. 1 illustrates that it varies with quantity: the peak is earlier in Q3 words than in Q2 words. So Fig. 8 shows the delay between the beginning of the syllable and the peak as a percentage of syllable duration.

It can be seen from Fig. 8 that the pitch peak was consistently later in Q2 words than in Q3 words irrespective of focus structure. In addition, the peak in the sentence-initial position appears to be somewhat earlier than in phrase-final words, again, irrespective of focus structure.

The generalised linear mixed models analysis showed that there was a weak three-way interaction between focus structure, phrase position and quantity with respect to peak timing ($\chi^2[2] = 8.9$, $p < .05$). Focus structure did not affect proportional peak timing. Position affected proportional peak timing such that the Q2 peak was later in phrase-initial words than in phrase-final words in the case of broad focus utterances ($p < .05$) and focus-final utterances ($p < .05$), but not focus-initial utterances. Similarly, the peak timing for Q3 was later in phrase-initial words than in phrase-final words in the case of broad and focus-final sentences ($p < .0001$), but not focus-initial sentences. Quantity affected peak timing in the phrase-initial words of broad focus, focus-final and focus-initial sentences (all ps < .0001) and in the phrase-final words of broad focus, focus-final and focus-initial sentences (all ps < .0001).

2.4.2. Duration

Although the study and experiment concentrated mainly on pitch prominence, data showing word duration as a function of phrase position, focus structure and word order are presented in Fig. 9. Fig. 9 shows that the duration of sentence-final object NPs is the longest. Most likely, this is due to the case-suffix *-t* (see Table 1 in Section 2.1.1). In addition, phrase-final words are lengthened in sentences in which the narrow focus is on the last word.

The generalised linear mixed effects analysis of the fixed effects of phrase position, focus position and word order on the dependent variable duration (in ms) showed an interaction between focus position and phrase position ($\chi^2[2] = 51.2$, $p < .0001$). Phrase position affected word duration in broad ($p < .0001$) and focus-final sentences ($p < .0001$) but not focus-initial sentences. There was a difference between final and broad focus sentences with respect to the duration of phrase-final words ($p < .05$) and an interaction between focus position and word order ($\chi^2[1] = 221.2$, $p < .0001$). Phrase position affected word duration in SVO ($p < .0001$) but not in OVS sentences. Word order affected the duration of phrase-initial and phrase-final words ($p < .0001$).

The effect of the interactions between quantity and focus structure and phrase position on word duration were also investigated (Fig. 10). Once again, phrase-final words had

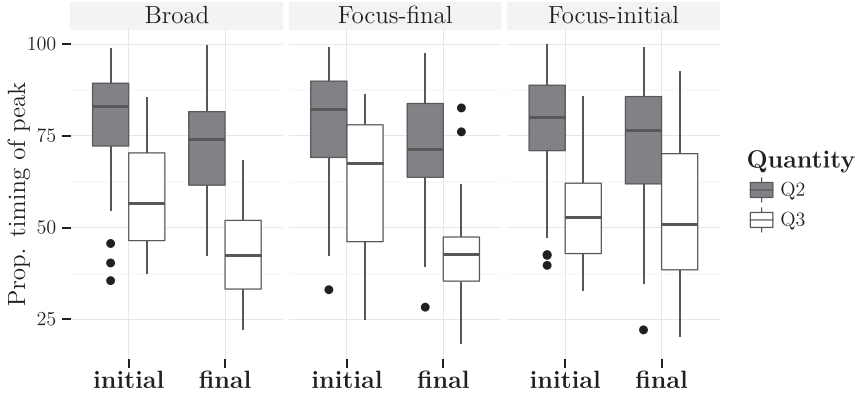

Fig. 8. Timing of pitch peak relative to the duration of the stressed syllable, shown as a function of focus position (broad; focus-initial; focus-final), phrase position (initial; final) and quantity (Q2; Q3).

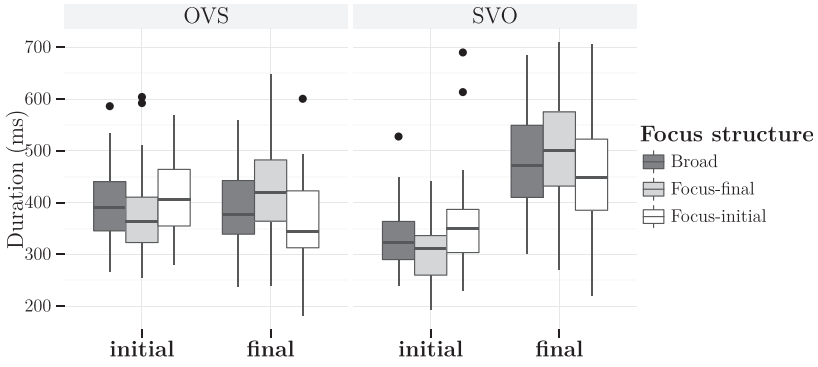

Fig. 9. Word duration as a function of focus structure (broad; focus-initial; focus-final), phrase position (initial; final) and word order (SVO; OVS).

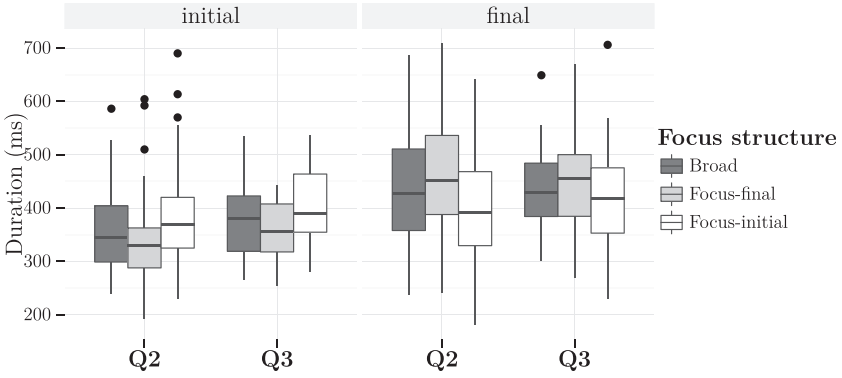

Fig. 10. Word duration as a function of focus structure (broad; focus-initial; focus-final), phrase position (initial; final) and quantity (Q2; Q3).

the longest duration, irrespective of quantity degree, and phrase-final words were longer in phrases with sentence-final narrow focus.

Another generalised linear mixed effects analysis with duration as the dependent variable and quantity, phrase position and focus structure as fixed effects showed that there was an interaction between phrase position and focus structure ($\chi^2[1] = 51.2$, $p < .0001$). *Post hoc* Tukey tests showed that there was a difference between initial focus and broad focus sentences with respect to duration of phrase-final words ($p < .05$) but not between broad and final focus sentences. There was no difference between broad, final and initial focus with respect to the duration of phrase-initial words. Phrase position affected word duration in broad and final focus sentences (both $ps < .0001$ respectively) but not initial focus sentences.

2.4.3. Declination

Global characterisations of intonation phrases were produced by fitting linear regression lines to the high (H) and to low points (L) of f0 excursions in individual utterances (see Fig. 11). The line through highs is called the topline and the line through lows is the baseline.

The lines in Fig. 12 show the average baselines and toplines against the background of the peaks (H) and valleys (L). It can be seen from Fig. 12 that the topline and baseline slopes show different trends: the topline slope interacts with the focus structure and word order of a sentence whereas the baseline slope does not. The baseline slope is slightly negative (ca −2 or −2.5) in all conditions. The topline slopes of focus-initial sentences are negative, varying from −7.3 in SVO sentences to −5.2 in OVS sentences.

Focus-final utterances have slightly positive topline slopes: 1.6 and 1.0 in SVO and OVS sentences respectively. Notably, in the broad focus condition, the two word orders have different toplines: in SVO sentences the topline is positive with a slope of 1.1, whereas in the OVS sentences it is negative with a slope of −2.1. The topline and baseline slopes cross in focus-initial utterances because the topline slopes are much steeper than the baseline slopes due to the post-focal compression.

A generalised linear mixed model with straight-line parameter slope as dependent variable and fixed effects of word order (SVO vs. OVS), focus structure (broad; focus-initial; focus-final) and type of regression line (topline; baseline) confirmed the above observations. There was a three-way interaction between word order, focus structure and regression type ($\chi^2[1] = 8.6$, $p < .05$). *Post hoc* Tukey tests on topline showed that word order affected the topline in broad focus sentences ($p < .0001$) but not in initial and final focus sentences. In SVO sentences there were differences between the toplines of initial and broad focus sentences ($p < .0001$) and initial and final focus sentences ($p < .0001$) but not broad and final focus sentences. In the case of OVS sentences all the pairwise comparisons were significant (all $ps < .0001$): final vs. broad focus; initial vs. broad focus and initial vs. final focus. Similar *post hoc* tests on baseline revealed no differences between conditions. There was a difference between the topline and baseline slope in broad focus SVO sentences ($p < .0001$), focus-final SVO sentences ($p < .0001$), focus-initial SVO

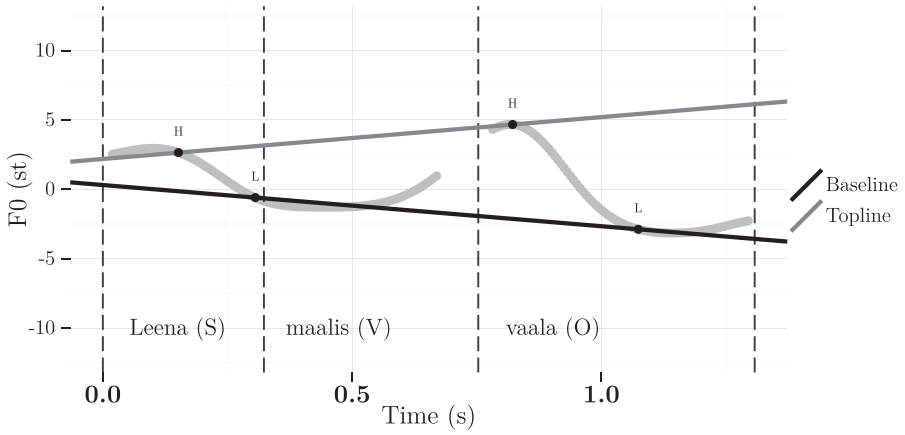

Fig. 11. Regression lines (dark grey and black indicate the topline and baseline respectively) fitted to the high ("H") and low ("L") points annotated on the f0 contour (light grey) for the utterance *Leena maalis vaala* (Leena drew a whale). Dashed vertical lines indicate word boundaries.

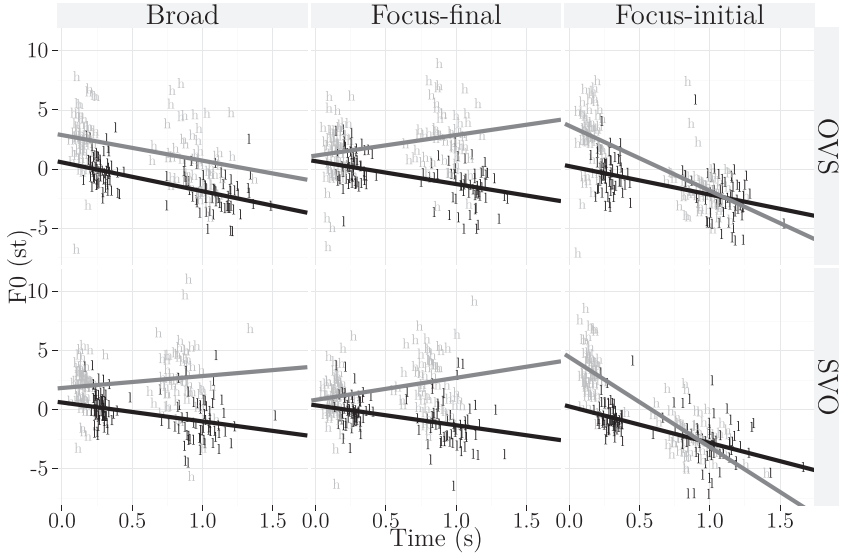

Fig. 12. Regression lines with slopes and intercepts aggregated over slopes of regression lines fitted to high ("h", see Fig. 3 for reference) or low points ("l", see Figs. 3 and 11 for reference) of excursions in every utterance. Plots in columns and rows show the mean regression lines (grey = topline, black = baseline) for focus position and word order respectively.

sentences ($p < .0001$), also focus-final OVS sentences ($p < .0001$) and focus-initial OVS sentences ($p < .0001$), but not in broad focus OVS sentences.

3. Discussion

The experiment investigated the effects of focus structure (broad; final; initial focus), word order (SVO; OVS) and phrase position (phrase-final; phrase-initial) on intonation of Estonian sentences. Intonation was quantified in terms of the range and slope of local f0 excursions on noun phrases and the topline and baseline slopes of the overall f0 trend. At word level, duration and peak timing were also examined. In the following, I first summarise the main results by dependent variable and then I relate the tonal variables to the two main research questions in separate subsections.

The trends in pitch range and slope corresponded and demonstrated that focus structure interacts strongly with the phrase position of a word. Local scaling of f0 showed that the peaks preceding the focal words and the peaks following the focal words were very narrow in pitch range. These results indicate that in Estonian narrow focus does affect the relative prominence of the accents in a phrase.

A related result was an effect of narrow focus that was highly dependent on the phrase position of the focal word. A comparison between broad focus and initial narrow focus showed that the f0 excursion on the first word was greater in the utterance where the first word was in focus than in the utterance without the narrow focus (broad focus sentence). This result is consistent with a well-known effect of narrow focus, namely that it causes an increase in the f0 range on words in sentence focus (e.g. Baumann et al., 2006; Breen et al., 2010; Chen & Braun, 2006; Cooper et al., 1985; Hanssen et al., 2008; Suomi et al., 2003; Xu & Xu, 2005).

However, the results also showed that relative to broad focus context, the pitch was not expanded on the last word in narrow focus. Notably, the broad and final focus sentences did not differ with respect to the pitch of the first word either. This result supports the earlier finding that, in Estonian, narrow focus does not affect sentence intonation (see e.g. Mihkla et al., 2015; Sahkai et al., 2013a) and demonstrates once again that sentence-final narrow focus does

not necessarily differ from broad focus in regard to prominence relations.

The experimental materials consisted of target words varying in word quantity. I did not control the quantity factor, as I assumed that the tonal feature of quantity (peak alignment) would not interact with phrase-level f0. The unbalanced representation of the quantity degrees notwithstanding, I also examined the effect of quantity on peak timing, size of f0 excursion (slope) and word duration. Although the results should be interpreted cautiously, they are consistent with predictions derived from the literature. The peak was later in Q2 words than in Q3 words, as predicted (Eek, 1974; Lehiste, 1960; Lippus et al., 2013; Mihkla & Kalvik, 2011). Word quantity interacted with focus structure to some extent, but generally the size of f0 excursions depended mainly on the presence or absence of sentence focus. As has been demonstrated previously, the three-level classification of quantity depends not on pitch range but on the timing of the pitch peak (see also Lippus et al., 2013).

In addition, pitch peak interacted with phrase position: the peak was earlier in phrase-final than in phrase-initial words. As argued elsewhere (Plüschke, 2013; Salveste, 2015), this might be due to the proximity of the intonational boundary. The effect of phrase position was less pronounced in utterances with a narrow focus at the beginning of the phrase. This result may well arise from the lack of a proper peak in the post-focal positions because, as was seen in the results, f0 excursions at the end of phrases with phrase-initial focus were very low and narrow in range.

The aim of the experiment was to examine the effect of focus structure on intonation. Therefore, the materials were not controlled for the examination of focus effect in duration. However, an evaluation of duration in relation to the same factors as f0 revealed two important findings. First, despite the highly variable segmental structure of the target words, I detected the effect of narrow focus on the duration of phrase-final word. This result corroborates earlier studies showing that words in narrow focus tend to be longer than words in broad focus (Sahkai et al., 2013a; Mihkla et al., 2015). Second, the results indicated that phrase-final position causes lengthening, if the sentence does not contain the narrow focus of phrase-initial word (for more on phrase-final lengthening in Estonian see Plüschke, 2013).

To provide a global analysis of f0, regression lines were fitted to local f0 maxima (high points in the high falling f0 excursion – peaks, see Fig. 3 for reference) and to local f0 minima (low points in the high falling f0 excursion – valleys). Word order and focus position affected topline slope. The topline was steeply negative for focus-initial utterances whereas it was slightly positive for focus-final utterances. The steep negative slope also indicates that the peak on the second noun phrase (NP) was scaled very low in comparison with the peak on the first NP. The positive slope, however, indicates that the peak on the second NP was scaled higher than the peak on the first NP. This indicates an *upstepped* pitch accent on the sentence-final objects in broad and final focus utterances. This result provides further evidence that in utterances with SVO word order, broad focus sentences and sentences with sentence-final narrow focus have similar global and local pitch characteristics.

In addition, the results for toplines show that in Estonian negative declination occurs in a specific context (initial focus) and where there is a specific word order (OVS) but otherwise, the topline is positive, which is not typical given the reports on other languages (see e.g. Liberman & Pierrehumbert, 1984; Swerts et al., 1996; Yuan & Liberman, 2014). These results are not easy to explain. First, it is possible that they are a side-effect of the experimental setting, as speakers knew that each utterance would be followed by another and might therefore have tended to use a list-like intonation pattern. However, based on qualitative evaluation during the annotation process, the utterances did not sound as if they were part of a list. Second, the syntactic structure of the target sentences might have caused the topline slope to be positive. It has already been shown that subject and object NPs have an influence on sentence intonation, hence they are more likely to carry a pitch accent (see e.g. Gussenhoven, 1983; Truckenbrodt & Darcy, 2010). The outcome of this experiment might indicate that the sentence-final word is more likely to carry an upstepped high pitch accent if it is in the object NP position. Further investigation of this possibility would involve comparing object NPs with some other constituent, such as a modifier. Third, the topline slopes may have been positive due to the preceding question.

The third suggestion is supported by looking at the materials in which declination has been observed previously. Liberman and Pierrehumbert (1984), for instance, observed declination in utterances listing instances of a categorical entity (e.g. different kinds of berries). Swerts et al. (1996) used excerpts of Swedish news telegrams that probably constituted coherent narrative texts. Yuan and Liberman (2014) drew their materials from Mandarin broadcast news speech corpora and it seems likely that these also constituted coherent narrative texts. It is possible that when answering questions speakers use upstepped pitch accents to try to attract the listener's attention to their answer; this might be why most of the utterances in this study showed an overall rise rather than declination in pitch. Based on the assumption that spontaneous speech is more interactive, a recent study by Asu, Lippus, Salveste, and Sahkai (2016) provided indirect evidence for this proposal. The authors intended to measure declination in spontaneous speech, but had to exclude almost 50% of the intonational phrases from evaluation. It appears, therefore, that declination tends to be less frequent in interactive speech and in Estonian.

The baseline was slightly negative for all utterance types and was not influenced by focus structure or word order. Whilst in SVO utterances baselines were mostly negative and toplines mostly positive and they differed in steepness (broad focus vs. focus-final), in utterances with OVS word order both lines had a negative slope. This indicates that OVS word order might be different from SVO word order with respect to semantics.

3.1. Pitch prominence as a function of focus position

The first goal of the experiment was to find support for the hypothesis that narrow focus affects the relative prominence or distribution of pitch peaks.

The location of narrow focus in the phrase strongly modulated the prominence relations between the pitch accents. The compression effects detected in words that were not in focus also indicate that broad focus was associated with a two-peak contour and narrow focus with a one-peak contour, which in turn demonstrates that focus structure affects the distribution of pitch accents within a phrase.

As suggested in Section 1.2, the distribution of pitch accents may be a much more important signal of sentence focus to the listener who is assessing the semantics of an utterance. This suggestion is supported by a study by Krahmer and Swerts (2001) that showed that Dutch listeners use the distribution of pitch accents and their relative prominence as cues to sentence focus, rather than the range of a single pitch accent on a word. I suggest that the distribution of pitch accents is also a psychoacoustically salient cue for Estonian listeners.

In addition to phrase-level f0 modification, I also found local f0 modification across the different conditions: the pitch range of the narrow focus of phrase-initial words was greater than its counterpart in broad focus sentences. Given that in these sentences the decreased number of pitch accents was already present as a very salient acoustic indicator of focus, expansion of the f0 excursion is, in one sense, a redundant acoustic cue to sentence-initial focus. Local modulation of pitch range should not, therefore, be entirely excluded as a potential linguistic indicator of focus in Estonian, but it is predicted that it does not occur consistently and is speaker-specific.

The regression lines fitted to points annotated on f0 excursions in semantically critical parts of utterances – at the focal and non-focal words – were affected by focus position. A strongly negative topline indicated that the phrase-initial word was in focus and that the initial focus was followed by a narrow-ranged pitch accent, whereas a slightly positive topline indicated that the phrase-final word was in focus and the pre-focal word might have carried a pitch accent (recall that the broad and final focus sentences did not differ with respect to the pitch of the first word). Thus topline slope can be used to quantify the relative prominence of pitch accents. Although this kind of application of regression lines is quite laborious (it requires manual annotation of f0 excursions), it allowed quantification of the post-focus f0 drop, mentioned by Cooper et al. (1985). The implication for research on declination is that the rough regression lines fitted to pitch contours might consist of more fine-grained information than is generally assumed.

3.2. Phrase-level f0 in interaction with word order

The second goal of the experiment was to explore whether intonation was affected more by word order or sentence focus?

Word order did not interact with narrow foci neither with respect to either the size of the f0 excursion or the overall f0 trend. Clearly, in utterances with both types of word order (SVO; OVS) focus was aligned with a pitch accent. This result is consistent with that of Sahkai et al. (2013b) and the model by Erelt et al. (1993) discussed in Section 1.2. The outcome provides further support to the earlier findings that phrase-level f0 does indeed have semantic implications that sit alongside pragmatically-driven variation in word order (e.g. Face & D'Imperio, 2005; Keller & Alexopoulou, 2001; Skopeteas et al., 2009; Välimaa-Blum, 1993; Vainio & Järvikivi, 2007).

Word order had an influence on pitch accent scaling and overall f0 trends in broad focus contexts. The topline for SVO was slightly positive, whereas the topline for OVS was more strongly negative. The range and slope of the phrase-final subject NP in broad focus utterances were greater than range and slope of the post-focal, phrase-final object NP (mean slope of 20 st/s compared with a mean slope of 6 st/s). This suggests that sentence-final subject NPs (OVS word order) carry a fully-fledged pitch accent. I infer from these values that the negative topline of OVS utterances reflects declination of two high pitch accents in sentences with OVS word order. The question then is what properties of this word order trigger a declination that was not present in SVO word order sentences.

My suggestion is that declination may reflect the presence of theme–rheme structure (or given-before or topic-comment structure) and cause the difference between SVO and OVS word order sentences with respect to prominence relations. In SVO utterances the phrase-final peak is higher than the phrase-initial peak, whereas in OVS utterances the phrase-initial peak is higher than the phrase-final peak.

Calhoun (2012) investigated the relative prominence of theme and rheme in New Zealand English and demonstrated that theme and rheme can be reliably distinguished on the basis of relative prominence: having the rheme at the end of the sentence causes an upstep of the subsequent pitch accent, whereas having the rheme at the beginning causes a decline in the subsequent pitch accent. Thus, the different topline trends in the Estonian speech production results may indicate that SVO word order reflects a theme–rheme structure whilst OVS word order, surprisingly, reflects a rheme–theme structure.

This is unexpected because the theoretical descriptions of Estonian word order claim that the variance in word order between SVO and OVS utterances is sensitive to theme–rheme structure such that the initial subject of a SVO utterance and the initial object of an OVS utterance both occupy the theme position (Erelt et al., 1993; see Section 1.2.2). However, the results from this experiment suggest that if the context does not specify relevant factors such as the location of the narrow focus, the word order itself may indicate whether the rheme is located at the end or beginning of an utterance and trigger the appropriate relative scaling of pitch accents.

Another possibility is that the phrase-final subject NP is low in prominence despite having an f0 excursion with a slope of about 20 st/s. This suggestion is based on linguistic analyses which found that sentence-initial object NPs are more emphatic than, for example, sentence-final object NPs (Lindström, 2006; Tael, 1988), or in my terminology, corrective. This semantic property of OVS word order may trigger the perception that the focus of the sentence is on the first word, despite the presence of a fairly conspicuous f0 excursion on the last word. For example, Vainio and Järvikivi (2006) showed that Finnish listeners perceived the sentence-final adverb as more prominent than the sentence-final object although both constituents carried an f0 excursion of same size.

The intonation pattern that is characteristic of OVS word order might reflect a narrow corrective focus on the

phrase-initial object NP or the rheme–theme structure. The semantics of narrow focus and rheme overlap and are difficult to tease apart. These theoretical arguments about the function of OVS word order and the interpretation of the associated intonation pattern definitely deserve further investigation via perception experiments. A more direct, data-driven explanation for the OVS word order results is that the declination of peaks and the smaller f0 excursion on the sentence-final subject NP were caused by the second quantity degree (Q2). Unfortunately this could not be tested because the effect of quantity was not controlled. Words in the subject NP position were always Q2 words whereas words in the object NP position were either Q2 or Q3 words. Quantity degree influenced, albeit weakly, the pitch range of f0 excursions in broad focus such that it was smaller on Q2 words than on Q3 words. As the experimental materials did not include subject NPs containing Q3 words the possibility that there was a word quantity effect cannot be excluded.

All in all, the results of the experiment provide support for the hypothesis that f0 interacts with sentence focus so as to cause low f0 and narrow-ranged f0 excursions in the post-focal part of a sentence and hence changes to the distribution of pitch accents in the corresponding phrase. The results are compatible with the two-dimensional theory of information structure (see Section 1.2) and with the proposal that word order encodes the theme–rheme distinction whereas intonation encodes sentence focus. As argued in Section 1.2, theme–rheme structure interacts with focus so that both parts can contain the narrow focus (Baumann & Kügler, 2015; Krifka, 2007; Steedman, 1998).

In the case of Estonian I adopt the view that word order is sensitive to the theme–rheme dimension: theme is placed before rheme and this is reflected in the OVS and SVO word orders. If these word orders are embedded in a context where they also contain focus, then the location of pitch accent varies as a function of focus position. Similar results have been reported in other unrelated free word order languages such as Georgian and Finnish (see e.g. Skopeteas et al., 2009; Välimaa-Blum, 1993; Vainio & Järvikivi, 2007, respectively). In those languages word order may also be sensitive to discourse-structural functions or pragmatic factors other than those of intonation.

4. Conclusion

This study explored the interactions between sentence focus, intonation and word order in the free word order language Estonian. The experiment showed that sentence focus is signalled by f0 such that the f0 range differs between focal and post-focal words. The word in focus carries a prominence-lending f0 excursion – i.e. pitch accent – whereas the word following the focus is compressed in duration and pitch – i.e. deaccented. The focus-driven expansion of f0 and lengthening of a word depend on its position within the phrase. These local indicators, may, however, be rather unsystematic and speaker-dependent. The results showed that pitch distribution (number of pitch accents in a phrase) and global characteristics of phrasal pitch (e.g. pitch change across the whole utterance) consistently indicate the location of sentence focus.

In conclusion, intonation, alongside word order, is an integral component of the meaning of a sentence in the free word order language Estonian. The implication for the general model of information structure is that as linguistic indicators of focus, word order and intonation are not mutually exclusive; in languages that exploit both they may encode different linguistic information.

Acknowledgments

This work was partly funded by Estonian Research Council Grant IUT2-37 and the project of The National Programme for Estonian Language Technology EKT71. The first draft of this paper gained a lot from valuable corrections and comments by Lia Saki Bučar Shigemori, Pärtel Lippus, Hanna Ruch and Eva Liina Asu. Furthermore, the author is very grateful to Jonathan Harrington for supervising her doctoral research at the University of Munich, part of which the paper covers. The author is extremely grateful for the suggestions and comments of the four anonymous reviewers of the paper.

References

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes, 23*, 495–527. http://dx.doi.org/10.1080/01690960801920099.

Asu, E. L. (2004). *The phonetics and phonology of Estonian intonation* (Ph.D. thesis). University of Cambridge.

Asu, E. L. (2005). Towards a phonological model of Estonian intonation. In M. Langemets & P. Penjam (Eds.), *Proceedings of the second baltic conference on human language technologies, Tallinn 4–5 May 2005* (pp. 95–100). Tallinn: Tallinn University of Technology and Institute of the Estonian Language.

Asu, E. L. (2006). Rising intonation in Estonian: An analysis of map task dialogues and spontaneous conversations. In R. Aulanko, L. Wahlberg, & M. Vainio (Eds.), *Fonetiikan Päivät 2006—The phonetics symposium 2006* (pp. 1–8). Helsinki: Helsinki University.

Asu, E. L., Lippus, P., Salveste, N., & Sahkai, H. (2016). F0 declination in spontaneous Estonian: Implications for pitch-related preplanning. *Proceedings of speech prosody, Boston 31 May–3 June 2016*. http://dx.doi.org/10.21437/SpeechProsody.2016-234.

Asu, E. L., & Nolan, F. (1999). The effect of intonation on pitch cues to the Estonian quantity contrast. In J. Ohala (Ed.), *Proceedings of the 14th international congress of phonetic sciences* (pp. 1873–1876). San Francisco: University of California.

Asu, E. L., & Nolan, F. (2007). The analysis of low accentuation in Estonian. *Language and Speech, 50*, 567–588.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. http://dx.doi.org/10.18637/jss.v067.i01.

Baumann, S. (2006). *The intonation of givenness—Evidence from German* (Ph.D. thesis). Saarland University.

Baumann, S., & Grice, M. (2006). The intonation of accessibility. *Journal of Pragmatics, 38*, 1636–1657.

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains—Categorical or gradient? In *Speech prosody—Proceedings of the third international conference* (pp. 301–304). Dresden: TUDPress.

Baumann, S., & Kügler, F. (2015). Prosody and information status in typological perspective—Introduction to the special issue. *Lingua, 165*, 179–182.

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [computer program]. URL: http://www.praat.org/.

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

Calhoun, S. (2012). The theme/rheme distinction: Accent type or relative prominence? *Journal of Phonetics, 40*, 329–349.

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.

Chen, Y., & Braun, B. (2006). Prosodic realization of information structure categories in Standard Chinese. In *Speech prosody—Proceedings of the third international conference*. Dresden: TUDPress.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse processes: Advances in research and theory* (pp. 1–40). Norwood, New Jersey: Ablex Publishing Corporation.

Cohen, A., Collier, R., & t'Hart, J. (1982). Declination: Construct or intrinsic feature of speech pitch? *Phonetica, 39*, 254–273.

Cole, J., Mo, Y., & Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes, 25*, 1141–1177.

Cooper, W., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *Journal of Acoustical Society of America, 77*, 2142–2155.

Cruttenden, A. (2006). The de-accenting of given information: A cognitive universal? In G. Bernini & M. L. Schwartz (Eds.), *Pragmatic organization of discourse in the language of Europe* (pp. 311–355). The Hague: Mouton de Gruyter.

Daneš, F. (1960). Sentence intonation from a functional point of view. *Word, 16*, 34–54. http://dx.doi.org/10.1080/00437956.1960.11659719.

Daneš, F. (1966). A three-level approach to syntax. In F. Daneš (Ed.), *Travaux linguistiques de Prague* (pp. 225–240). Prague: Academia editions de L'Académie Tchécoslovaque des Sciences.

D'Imperio, M. (2000). *The role of perception in defining tonal targets and their alignment* (Ph.D. thesis). The Ohio State University.

Eek, A. (1974). Observations on the duration of some word structures. In A. Eek (Ed.), *Estonian papers in phonetics* (pp. 31–57). Tallinn: Academy of Sciences of the Estonian S.S.R..

Eek, A. (1980). Estonian quantity: Notes on the perception of duration. In A. Eek (Ed.), *Estonian papers in phonetics* (pp. 5–30). Tallinn: Academy of Sciences of the Estonian S.S.R.

Eek, A. (1983). Kvantiteet ja rõhk eesti keeles (I). (Estonian) [Quantity and accent in Estonian]. *Keel ja Kirjandus [(Journal of) Language and Literature], 9*, 481–489.

Ehala, M. (2006). The word order of Estonian: Implications to universal language. *Journal of Universal Language, 7*, 49–89.

Erelt, M., Kasik, R., Metslang, H., Rajandi, H., Ross, K., Saari, H., ... Vare, S. (1993). *Eesti keele grammatika. Süntaks*. Tallinn: Eesti Teaduste Akadeemia. Keele ja Kirjanduse Instituut.

Face, T. L., & D'Imperio, M. (2005). Reconsidering a focal typology: Evidence from Spanish and Italian. *Rivista di Linguistica, 17*, 271–289.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics, 36*, 680–703.

Firbas, J. (1966). On defining the theme in functional sentence analysis. In F. Daneš (Ed.), *Travaux Linguistiques de Prague* (pp. 267–280). Prague: Academia editions de L'Académie Tchécoslovaque des Sciences.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of Acoustical Society of America, 27*, 765–768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126–152.

Gryllia, S. (2009). *On the nature of preverbal Focus in Greek. A theoretical and experimental approach* (Ph.D. thesis). University of Leiden.

Gundel, J. (1985). 'Shared knowledge' and topicality. *Journal of Pragmatics, 9*, 83–107.

Gussenhoven, C. (1983). Testing the reality of focus domains. *Language and Speech, 26*, 61–80.

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* (pp. 83–100). Dordrecht: Springer, Netherlands.

Gussenhoven, C., Repp, B., Rietveld, A., Rump, W., & Terken, J. (1997). The perceptual prominence of fundamental frequency peaks. *Journal of Acoustical Society of America, 102*, 3009–3022.

Halliday, M. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics, 3*, 199–244.

Hanssen, J., Peters, J., & Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. In P. A. Barbarosa, S. Madureira, & C. Reis (Eds.), *Speech prosody—Proceedings of the fourth conference* (pp. 609–612). Brazil, Campinas.

Keller, F., & Alexopoulou, T. (2001). Phonology competes with syntax: Experimental evidence for the interaction of word order and accent placement in the realization of information structure. *Cognition, 79*, 301–372.

Kiss, K. É. (1995). Introduction. In K. É. Kiss (Ed.), *Discourse configurational languages. Oxford studies in comparative syntax* (pp. 3–27). New York, Oxford: Oxford University Press.

Knight, R.-A. (2003). *Peaks and Plateaux: The production and perception of high intonational targets in English* (Ph.D. thesis). University of Cambridge.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of Acoustical Society of America, 118*, 1038–1054.

Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication, 34*, 391–405.

Krifka, M. (2007). Basic notions of information structure. In C. Fery & M. Krifka (Eds.), *Interdisciplinary studies on information structure* (Vol. 6, pp. 13–56). Potsdam: Universitätsverlag.

Ladd, D. R. (2008). Intonational phonology. In *Cambridge studies in linguistics* (Vol. 119). Cambridge: Cambridge University Press.

Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Lehiste, I. (1960). Segmental and syllabic quantity in Estonian. *American Studies in Uralic Linguistics, 1*, 21–82.

Lehiste, I. (1997). Search for phonetic correlates in Estonian prosody. In I. Lehiste & J. Ross (Eds.), *Estonian prosody: Papers from a symposium* (pp. 11–35). Tallinn: Institute of Estonian Language.

Lenerz, J. (1977). *Zur Abfolge nominaler Satzglieder im Deutschen. (German) [About the order of nominal constituents in German]*. Tübingen: Gunter Narr.

Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure. studies in phonology. Presented to Morris Halle by his teacher and students* (pp. 155–233). Cambridge, MA: The Massachusetts Institute of Technology.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of Acoustical Society of America, 32*, 451–454.

Lindström, L. (2002). Veel kord subjekti ja predikaadi vastastikusest asendist laiendi järel. (Estonian) [More on positions of subject and predicate located after a modifier]. In M. Erelt (Ed.), *Emakeele seltsi aastaraamat* (pp. 87–105). Tallinn: Eesti Teaduste Akadeemia.

Lindström, L. (2004). Sõnajärg lause tuumaargumentide eristajana eesti keeles. In L. Lindström (Ed.), *Lauseliikmeist eesti keeles. Tartu Ülikooli eesti keele õppetooli preprindid* (pp. 40–49). Tartu.

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

Lippus, P., Asu, E. L., Teras, P., & Tuisk, T. (2013). Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian. *Journal of Phonetics, 41*, 17–28.

Lippus, P., Pajusalu, K., & Allik, J. (2009). The tonal component of Estonian quantity in native and non-native perception. *Journal of Phonetics, 37*, 388–396.

Lippus, P., Pajusalu, K., & Allik, J. (2011). The role of pitch cue in the perception of the Estonian long quantity. In S. Frota, G. Elordieta, & P. Prieto (Eds.), *Prosodic categories: Production, perception and comprehension* (pp. 231–242). Dordrecht: Springer.

Mihkla, M., & Kalvik, M.-L. (2011). Significant features in Estonian word prosody. In W.-S. Lee & E. Lee (Eds.), *Proceedings of XVIIth international conference of phonetic sciences, Hong Kong 17–21 August*. .

Mihkla, M., Sahkai, H., & Kalvik, M.-L. (2015). Acoustic correlates of emphasis in Estonian. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th international congress of phonetic sciences, Glasgow 10–14 August*. .

Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *Journal of Acoustical Society of America, 66*, 363–369.

Plüschke, M. (2013). *Peak alignment in Estonian* (Ph.D. thesis). Ludwig-Maximilians-Universität München.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). New York: Academic Press.

R Core Team (2017). R: A language and environment for statistical computing. URL: http://www.R-project.org/.

Rathcke, T., & Harrington, J. (2006). Is there a distinction between H+!H* and H+L* in standard German? Evidence from an acoustic and auditory analysis. In *Speech prosody—Proceedings of the third international conference*. Dresden: TUDPress.

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica, 27*, 53–94.

Remmel, M. (1963). Sõnajärjestus eesti lauses. Deskriptiivne käsitlus. (Estonian) [Word order of Estonian sentences. A descriptive survey]. In R. Kress (Ed.), *Eesti keele süntaksi küsimusi* (pp. 216–217). Tallinn: Eesti Riiklik Kirjastus.

Robinson, D., & Dadson, R. (1956). A redetermination of the equal-loudness relations for pure tones. *British Journal of Applied Physics, 7*, 166–181.

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

Sahkai, H., Kalvik, M.-L., & Mihkla, M. (2013a). Prosody of contrastive focus in Estonian. In *Proceedings of the 14th annual conference of the international speech communication association, 25–29 August* (pp. 315–319). Lyon, France: International Speech Communication Association.

Sahkai, H., Kalvik, M.-L., & Mihkla, M. (2013b). Prosodic effects of the information structure in Estonian. In E. L. Asu & P. Lippus (Eds.), *Nordic prosody—Proceedings of the XIth conference, Tartu 2012* (pp. 323–332). Frankfurt am Main: Peter Lang Verlag.

Salveste, N. (2010). Variation of pitch cues in the perception of Estonian quantities. In S. Werner & T. Kinnunen (Eds.), *Proceedings of the XXVI phonetics Fonetiikan Päivät—The phonetics symposium* (pp. 42–47). Finland, Mekrijärve: University of Eastern Finland.

Salveste, N. (2015). *On the pragmatic and semantic functions of Estonian sentence prosody* (Ph.D. thesis). Ludwig-Maximilians-Universität München.

Sgall, P., Hajicová, E., & Panevová, J. (1986). *The meaning of the sentence in its semantic and pragmatic aspects*. Prague: D. Reidel Publishing Company.

Siptár, P., & Terköczy, M. (2000). *The phonology of Hungarian*. Oxford: Oxford University Press.

Skopeteas, S., Féry, C., & Asatiani, R. (2009). Word order and intonation in Georgian. *Lingua, 119*, 102–127.

Steedman, M. (1998). Information structure and the syntax-phonology interface. In *IRCS technical reports series institute for research in cognitive science*. University of Pennsylvania. URL: http://repository.upenn.edu/ircsreports/49.

Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry, 31*, 649–689.

Steedman, M. (2014). The surface-compositional semantics of English intonation. *Language, 90*, 2–57.

Suomi, K., Toivanen, J., & Ylitalo, R. (2003). Durational and tonal correlates of accent in Finnish. *Journal of Phonetics, 31*, 113–138.

Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics, 30*, 629–654.

Swerts, M., Strangert, E., & Heldner, M. (1996). F0 declination in read-aloud and spontaneous speech. In *Proceedings of the fourth international conference on spoken language processing, Philadelphia, PA 3–6 October* (Vol. 3, pp. 1501–1504).

Szendröi, K. (2003). A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review, 20*, 37–78.

Tael, K. (1988). *Sõnajärjemallid eesti keeles (võrrelduna soome keelega). (Estonian) [Word order patterns in Estonian in comparison with Finnish]*. Technical Report Eesti NSV Teaduste Akadeemia Keele ja Kirjanduse Instituut Tallinn.

Truckenbrodt, H., & Darcy, I. (2010). Object clauses, movement and phrasal stress. In N. Erteschik-Shir & L. Rochman (Eds.), *The sound patterns of syntax* (pp. 189–216). Oxford: Oxford University Press.
Turk, A. E., & Sawusch, J. R. (1996). The processing of duration and intensity cues to prominence. *Journal of Acoustical Society of America, 99*, 3782–3790.
Vainio, M., & Järvikivi, J. (2006). Tonal features, intensity, and word order in the perception of prominence. *Journal of Phonetics, 34*, 319–342.
Vainio, M., & Järvikivi, J. (2007). Focus in production: Tonal shape, intensity, and word order. *Journal of the Acoustical Society of America, Easy Letters, 121*, 55–61.
Välimaa-Blum, R. M. (1993). Intonation: A distinctive parameter in grammatical constructions. *Phonetica, 50*, 124–137.
Vilkuna, M. (1998). Word order in European Uralic. In A. Siewierska (Ed.), *Constituent order in the languages of Europe* (pp. 173–233). Berlin: Mouton de Gruyter.
Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics, 33*, 159–197.
Yuan, J., & Liberman, M. (2014). F0 declination in English and Mandarin broadcast news speech. *Speech Communication, 65*, 67–74.