Testing a Model of Estonian Intonation

Eva Liina Asu and Francis Nolan
University of Cambridge, UK
E-mail: ela21@cus.cam.ac.uk, fjn1@cus.cam.ac.uk

ABSTRACT

In this paper, a descriptive model of Estonian intonational phonology is presented, based on experimental investigation using controlled read materials. This model is tested on a corpus of task-oriented spontaneous speech. Characteristics found on the basis of read materials included a preference for peaked pitch accents which can be labelled as H*+L; the use of some nuclear rises (L*+H); and the occurrence, in statements only, of an H+L* pitch accent in nuclear position, with optional spreading to earlier positions. Phonological analysis of the spontaneous data reveals that the characteristics summarised in the model do apply to spontaneous speech. New phenomena include a possible distinction between two kinds of rising nucleus; the occurrence of a stepping pattern over the intonational phrase; and the confirmation of distributional regularities within intonational phrases with an H+L* nuclear pitch accent.

1. INTRODUCTION

Previous research on Estonian intonation [e.g. 1, 2, 3] has focused on read speech. On the basis of the analysis of this kind of data, it has been possible to draw up a preliminary descriptive model of Estonian intonation. The aim of the present paper is to test the phonological inventory of this model against the phonology of task-oriented spontaneous speech as elicited by the so-called 'map task'.

In the first autosegmental-metrical description of Estonian intonational phonology [4], based on an analysis of a read passage, a preliminary inventory consisting of five intonational pitch accents was established. The most frequently occurring pitch accent in different types of utterances both in nuclear position (the last pitch accent of the intonational phrase) and in prenuclear position is H*+L: a relatively sharply peaked accent where an accented high syllable (optionally preceded by a lower unaccented syllable) is immediately followed by a fall to a low tone. Arguably distinct is an ^H*+L (an upstepped variant of the fall) where the accented syllable is perceptually higher in pitch than the preceding accented syllable. It was necessary to distinguish between these two falls in nuclear position because of their functional characteristics. As shown in [3], statements and questions are differentiated by the height of a nuclear fall, and, furthermore, as becomes evident in [1], different types of yes/no question in Estonian can be partially characterised by the height of the nuclear pitch accent (in addition to different prenuclear downtrends, as explained below).

A pitch accent that occurred fairly frequently in the read passage used for establishing the preliminary inventory in [4], and that had not been previously described for Estonian, was a low level nucleus. This was initially labelled as a downstepped high tone followed by a low trailing tone (!H*+L). In [2], however, this pitch accent was re-analysed as a low target preceded by a high leading tone (H+L*). It was observed [4] that such low level nuclei never occur in questions but were common in statements, and that low accentuation can spread to prenuclear accents in an utterance where the nucleus is an H+L* [2]. It might be the case that this pitch accent is mainly characteristic of read speech and therefore an analysis of spontaneous speech data will shed more light on its distribution.

Traditionally rises have not been considered part of Estonian intonational inventory. The provisional inventory [4], however, included rises (L*+H) which were found in the nuclear position of certain questions and as a continuation marker. Although the fairly limited number of rises in the read passage confirms that a rise is not a common phonological choice, it is still one of the options available. Rises did not appear to involve an H% boundary tone; when sufficient post-nuclear material is available, a level plateau emerged, justifying the use of 0% (see 2.2 below). No evidence was found for an L boundary tone. Table 1 summarises the provisional tonal inventory in a format which follows that of [5].

Intonation events	Pitch description	Phonological analysis
Pitch Accents	Simple high	H*
	Simple low	L*
	Falling	H*+L
	Upstepped falling	^H*+L
	Low with high onset	H+L*
	Rising	L*+H
Boundary Tones	No pitch movement in the vicinity of the boundary	0%

Table 1: Provisional inventory of Estonian pitch accents and boundary tones based on read speech.

Additionally, from the functional point of view, different types of yes/no question in Estonian were studied in [1]. The results show that there exists a trading relationship

ISBN 1-876346-48-5

between morpho-syntactic marking and intonation: morpho-syntactically unmarked questions, and those where the morpho-syntactic marking is at the end of the sentence rather than the beginning, employ intonational marking in terms of a shallower downtrend over prenuclear accents.

2. METHOD

2.1 MATERIALS

Spontaneous speech was elicited with the help of a map task where one speaker has to describe a route on their map to another speaker. For this purpose two maps were designed where objects and their location were slightly varied so as to create discussion. The general principles of map task design set out in [6] were followed. The corpus of spontaneous speech for the present paper consists of recordings of six female speakers of Standard Estonian who were recorded doing the map tasks in pairs. Recordings were carried out in a quiet environment in Tartu, Estonia, using a Sony TCD D8 portable DAT recorder and two microphones attached to the recorder via an amplifier. The map task corpus contains about 30 min of dialogue.

2.2 PHONOLOGICAL ANALYSIS

The speech data was digitised at 16 kHz on a Silicon Graphics Unix workstation and an F0 contour was computed for all utterances using Xwaves+. All data was transcribed orthographically and labelled prosodically. Analysis was carried out with the labelling tool Labeler, developed in conjunction with the ToBI labelling conventions [7]. The phonological inventory used was primarily the IViE inventory for prosodic labelling [8], which is based on ToBI.

Three of the characteristics which differentiate the IViE labelling system from ToBI are that it allows a 'zero' boundary tone when there is no pitch change in the vicinity of an IP boundary, it treats nuclear falls as H*+L rather than as the combination H* L- L% (L- being a 'phrase accent', dropped in the IViE system), and it allows only 'left-headed' pitch accents (such as H*+L) to the exclusion of 'right-headed' pitch accents such as H+L*. The intonation patterns which emerged from [1, 3, 4] were compatible with the first two of these precepts of IViE, but the optimal analysis of the 'downstepped' nucleus mentioned above was shown to be H+L* [2], which requires the abandonment of the strict IViE constraint against right-headed pitch accents.

RESULTS AND DISCUSSION

3.1 GENERAL OBSERVATIONS

The nature of the speech elicited by the map task differs considerably from either the read story in [3, 4] or the carefully controlled read sentences of [1, 2]. The map task speech is consistently interactional, and is made up of mainly very short speaking turns, including feedback expressions. Since the participants are engaged in the collaborative and cognitively challenging task of agreeing a route on a map which, unknown to them at the start of the interaction, differs in crucial details, their speech is characterised by elliptical utterances, hesitations, requests for confirmation, overlapping speech turns, and fillers.

3.2 THE INVENTORY OF TONES

The map task speech, even though it diverges stylistically from the read speech analysed previously, exhibits all of the pitch accents listed in Table 1. In read speech H*+L is the 'default' pitch accent in both nuclear and pre-nuclear positions, and this is the case too in the map task recordings. However, its purely statistical dominance is challenged by the large number of rises found, particularly on feedback expressions such as *jaa/jah* ('yes'), *soo/nii/noo* ('so'), etc. Whilst these rises are hardly surprising from the perspective of many other languages, including English, they do fly in the face of the belief held by traditional accounts of Estonian intonation [e.g. 9] that rises are foreign to Estonian and extremely uncommon. The phonological analysis of the rises will be the subject of comment below in subsection 3.3.

Perhaps unexpectedly, low level accents preceded by a high unaccented syllable (H+L*) were widespread in the map task data. We had suspected that these might be a characteristic of the highly planned prosody associated with read statements. However, H+L* accents occurred in the speech of all six subjects. An example of H+L* on the nucleus is shown in Figure 1.

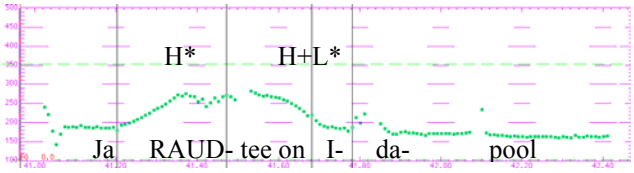

Figure 1: An example of an utterance with a low level nucleus (H+L*) ('And the railway is in the east').

The default pattern on feedback expressions was that shown in Figure 2. Although the predominant percept to which such utterances gives rise is a 'low rise', it is clear that these is also a marked downward onglide to the low accent. For such utterances we propose the analysis H+L* H%. This introduces a high boundary tone which was not found necessary in the analysis of the read data, where any nuclear rises could be accommodated as L*+H 0%.

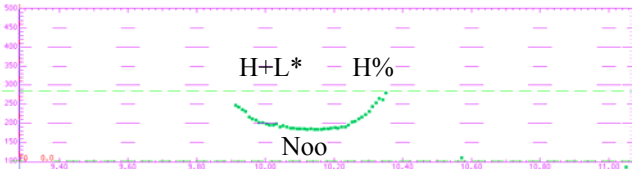

Figure 2: An example of a monosyllabic feedback expression with H+L* H% ('So').

3.3 TWO TYPES OF RISE

Independent support for adding H% to the inventory is found elsewhere in the map task data. Whilst the rises found all start low (i.e. are L* rather than H*), they appear, provisionally at least, to fall into two types according to whether the rise is implemented immediately or delayed until the boundary, as shown respectively in Figures 3 and 4. For these we adopt the analyses L*+H 0% and L* H%. Whilst phonetically this distinction is neutralised on short material, as in monosyllabic feedback expressions, we would suggest that H% has a specifically co-operative function in interaction (cf. the high-range rise and fall-rise in [5]), and this justifies the use of H+L* H% in the previous section for feedback expressions (see Figure 2).

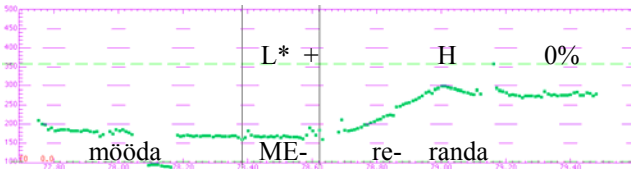

Figure 3: An example of an L*+H 0% (‘Along the sea coast’).

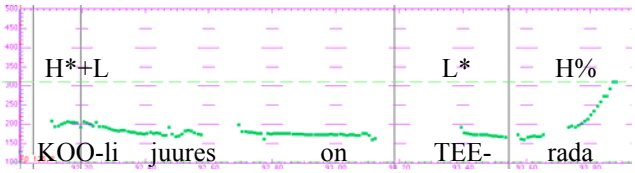

Figure 4: An example of an L* H% (‘There’s a path near the school’).

3.4 STEPPING PATTERN

In the map task data there occurs a pattern consisting of a sequence of steps between plateaus, like the traditional ‘stepping head’ of [10]. This pattern, which is illustrated in Figure 5, was not found in the read sentences previously analysed. For its analysis we adopt the ‘downstepping’ H+!H* pitch accent from ToBI, thereby adding a pitch accent to the inventory of Table 1.

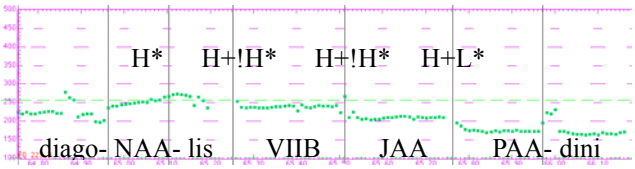

Figure 5: An example of the stepping pattern (‘Diagonally takes yes to the boat’).

A potential problem concerns the analysis of the last, nuclear, pitch accent, which in Figure 5 is transcribed as H+L*, even though it involves a ‘step’ rather similar to the preceding ones. In fact there seems to be a difference between stepping patterns which, as here, reach the speaker’s lowest pitch, and others found in the map task data which end on a non-low, ‘sustentive’ pitch. In Figure 5 the speaker’s final nuclear (H+)L* is at around 170 Hz. In Figure 6, which exemplifies the sustentive pattern for the same speaker, the pitch at the equivalent point is around 200 Hz. In the latter kind of stepping pattern the nuclear pitch accent could be analysed as H+!H*, the same as the pre-nuclear accents.

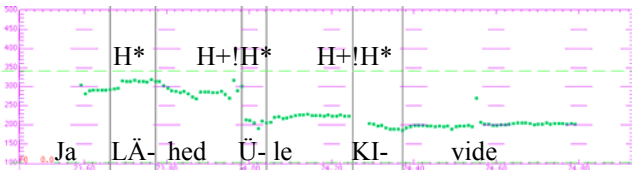

Figure 6: An example of the stepping pattern ending non-low (‘And you go across the stones’).

3.5 INTONATIONAL GRAMMAR

The H+L* accent, illustrated in Figure 1 with an H* pre-nuclear accent, is frequently preceded instead by one or more low accents. The maximal case is shown in Figure 7, where all the accents are (H+)L*.

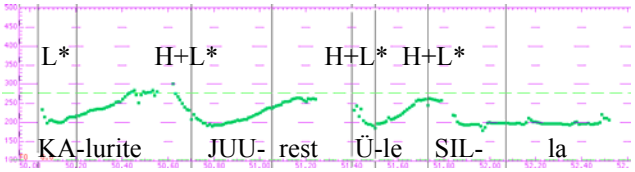

Figure 7: An example of an utterance with a succession of H+L* accents (‘Near the fishermen over the bridge’).

The interesting regularity which emerges from both the map task data and the previously analysed read data is that the choice of a low pitch accent in position N is correlated with H+L* being chosen in position N+1. That is, a final, nuclear H+L* opens up the possibility of (H+)L* for the penultimate accent; and only if the penultimate accent is H+L* can the antepenultimate pitch accent be low. Thus the sequence H*+L, H*, H+L*, H+L* is well formed, but L*, H+L*, H*, H+L* is not. This means that the selection of pitch accents cannot be accounted for by a finite state grammar in which all pitch accents are available independent of other choices as proposed in [11]. At the very least, the choice of a low pitch accent at position N would have to restrict the choice of pitch accent at N+1 to a low pitch accent. This ‘left-to-right’ approach, however, seems to us to be intuitively incorrect in that it attributes a major choice defining the intonational character of the IP to an arbitrary pitch accent which, though sometimes the nucleus when there is only one low accent, may be a pitch accent of relatively little informational salience embedded in the pre-nuclear string. The formal solution to this problem lies outside the scope of this paper, but will depend on the recognition of a hierarchical structure to the IP, as proposed for instance by [12], and a mechanism to allow choices in the ‘(syntactic) head’ of the IP, i.e. the ‘nucleus’, to influence pre-nuclear positions.

4. CONCLUSIONS

The phonological inventory of spontaneous Estonian speech has been shown to be similar to that of read speech. All the pitch accents found previously in read speech have

ISBN 1-876346-48-5

been observed in the map task data; and, in the case of the H+L* pitch accent, the same distributional regularity within the IP has been confirmed.

But there are differences. Perhaps because of the interactional nature of the map task, evidence emerged for an H% boundary tone, whereas rises in the read data appear to be L*+H 0% (as evidenced by a high plateau which emerges when there is sufficient phonetic material). Further analysis will be required to confirm this. Additionally a stepping pattern, necessitating the use of H+!H*, was encountered for the first time in the spontaneous data.

Nonetheless the degree of correspondence between the two types of data is reassuring from the methodological point of view: read speech, necessary for controlled quantitative experiments, did not produce artefactual intonational phonology, and revealed a high proportion of the intonational phenomena of spontaneous speech.

REFERENCES

[1] E. L. Asu, "Downtrends in different types of question in Estonian", in *Proceedings of the Speech Prosody 2002 conference, 11-13 April 2002*, B. Bel and I. Marlien, Ed., pp. 143-146. Aix-en-Provence: Laboratoire Parole et Langage, 2002.

[2] E. L. Asu, "Intonational phonology of Estonian statements with low level nuclei", *Speech, Music and Hearing, TMH-QPSR, Vol. 44, Proceedings from Fonetik 2002,* pp. 165-168, 2002.

[3] E. L. Asu and F. Nolan, "The interaction of intonation and quantity in Estonian: an analysis of nuclear falls in statements and questions", in *Nordic Prosody. Proceedings of the VIII Conference, Trondheim 2000*. W. A. van Dommelen and T. Fretheim, Ed., pp. 23-32. Frankfurt am Main: Peter Lang, 2001.

[4] E. L. Asu, "An autosegmental-metrical analysis of Estonian intonation: preliminary results", *Proceedings of the Ninth International Congress for Fenno-Ugric Studies*, vol. 4, pp. 83-89, 2001.

[5] J. Fletcher, R. Wales, L. Stirling and I. Mushin, "A dialogue act analysis of rises in Australian English map task dialogues", in *Proceedings of the Speech Prosody 2002 conference, 11-13 April 2002*, B. Bel and I. Marlien, Ed., pp. 299-302. Aix-en-Provence: Laboratoire Parole et Langage, 2002.

[6] A. Anderson, et al., "The HCRC Map Task Corpus", *Language and Speech*, vol. 34 (4), pp. 351-366, 1991.

[7] http://www.ling.ohio-state.edu/~tobi/

[8] E. Grabe, *The IViE Labelling Guide*, www.phon.ox.ac.uk/~esther/ivyweb/guide.html

[9] P. Ariste, *Eesti Keele Foneetika*, Tartu: Tartu Riiklik Ülikool, 1977.

[10] J.D. O'Connor and G.F. Arnold, *Intonation of Colloquial English* (2nd edition), London: Longman, 1973.

[11] J.B. Pierrehumbert, *The Phonology and Phonetics of English Intonation*, unpublished doctoral dissertation, Cambridge MA: MIT, 1980.

[12] D.R. Ladd, *Intonational Phonology*, Cambridge: CUP, 1996.

ISBN 1-876346-48-5