Intonation of Finnish Verbs

Anja Arnhold[1,2], *Martti Vainio*[2], *Antti Suni*[2], *Juhani Järvikivi*[3]

[1] Department of Linguistics, University of Potsdam, Germany
[2] Department of Speech Sciences, University of Helsinki, Finland
[3] Max Planck Institute for Psycholinguistics, The Netherlands

anja.arnhold@gmail.com, martti.vainio@helsinki.fi, antti.suni@helsinki.fi, juhani.jarvikivi@mpi.nl

Abstract

A production experiment investigated the tonal shape of Finnish finite verbs in transitive sentences without narrow focus. Traditional descriptions of Finnish stating that non-focused finite verbs do not receive accents were only partly supported. Verbs were found to have a consistently smaller pitch range than words in other word classes, but their pitch contours were neither flat nor explainable by pure interpolation.

Index Terms: Finnish, verbs, intonation, accent

1. Introduction

It is commonly assumed that, in contrast to other word classes, Finnish finite verbs would be unaccented unless in narrow focus (e.g. [1], [2]); or that verbs receive weaker accents than the other words in a sentence (Peltonen (1901) as reported in [3]). However, there has been no study so far specifically designed to investigate whether this is the case; furthermore, most characterizations in the literature are rather impressionistic and not based on quantitative data.

Välimaa-Blum notes that the descriptions of Finnish verbs are reminiscent of Schmerling's ([4]) observation that English and German verbs are assigned a lower accent than their arguments ([1]:84). There has been an extensive debate on how this phenomenon should be accounted for. Against early purely syntactic explanations (e.g. [5]), Bolinger ([6]) has claimed that accent distribution is only a matter of the speaker's choice of focus. He thus assumes that finite verbs are not structurally different from other word classes, and that in apparent counterexamples, the verb is 'semantically poorer' than its accented argument. Since [7], it has usually been maintained that while information structure is important, there also is a structural difference between predicates and their arguments. On a whole, however, the discussion has largely focused on a few West Germanic languages (for a recent proposal see [8]). Reviewing data from a few more languages, Ladd ([9]:193) tentatively concludes that verbs are treated distinctly in some languages, while the difference is not made in others.

The present article takes a step towards further broadening the perspective. It reports a production study designed to assess the previous claims regarding Finnish finite verbs. In particular, we aimed at determining whether the pitch contours of verbs would indeed be special and particular to that word class. In other words, the question is whether the accentuation of verbs is only a matter of semantic content or free choice or whether there is a structural difference between verbs and their arguments. A more specific question was whether the pitch contours of finite verbs would be the result of pure interpolation, as would be expected if verbs do not receive accents. To assess these questions, we compared the contours of verbs in different positions with both each other and with those of adverbs in the same linear position within a sentence. To ensure that any possible differences were indeed structural, we concentrated on sentences without narrow focus and chose combinations of verbs and objects that would not make one predictable from the other, thus avoiding 'semantically empty' verbs (or objects). We investigated only transitive sentences, because the accent assignment for intransitive verbs is likely to depend on their semantic type (e.g. [8]).

2. Methods

2.1. Materials

The materials recorded for this study contained twelve different verbs in third person present tense embedded in simple transitive sentences – four for each of the types CVV, CV.CVV and CVV.CVV. While in order to facilitate a broad focus interpretation the relative order of the verbs and their arguments was always unmarked (SVO), the linear position of the finite verb was varied by the presence or absence of adverbials. The sentences contained either (a) an adverbial in preverbal position, (b) an adverbial in postverbal position, (c) a pre- and a postverbal adverbial, or (d) no adverbials.

Within each of the adverbial positions, syllable number was kept identical and syllable structure very similar. To provide a strong case of comparison for the supposedly unaccented verbs, most of the adverbials were chosen to be as semantically empty as possible, so that they could also be expected to be unaccented. This was the case for all the three adverbials in the position preceding the verb, whereas postverbally, two semantically empty adverbs, as well as two nouns in adessive case appeared.

The syllable structure was always the same for the subjects (CVV.CV) and objects (CV.CVC); there were 36 different lexical items in both cases.

Altogether, 72 sentences were recorded with each participant.

2.2. Speakers and recording procedure

Twenty-five native speakers of Finnish (18 female) took part in the experiment. The participants read the sentences as answers to prerecorded questions like *Mitä tapahtuu?* "What happens?" or *Mitä Olli sanoo?* "What does Olli say?", which induced broad foci in the answers. To keep the participants alert, the target sentences were interspersed with 72 filler sentences that were to be uttered with different kinds of focus structures according to the questions. There were five different pseudo-randomized orders of the fillers and experimental sentences, each of which was used for five informants. The

informants were instructed to speak lively and answer as if their responses were part of an interesting conversation.

The recordings took place in a sound-proof room at the Department of Speech Sciences of the University of Helsinki. The prompts were played over loudspeakers with pauses of sufficient length to allow for the participants' answers. The speakers' responses were recorded directly to a computer hard-disk using a high-quality condenser microphone. A high-quality analogue-to-digital converter was used with 16 bit quantization and 44.1 kHz sampling frequency.

2.3. Editing and analyses

Responses including slips of the tongue or hesitations (77 cases), 103 sentences not spoken as one prosodic phrase, as well as 58 sentences with colloquial shortening of the adverbials were discarded from the analyses. Furthermore, eight responses were lost due to technical problems. Altogether, 1554 sentences were retained for the analyses.

For the acoustic measurements, the program Praat ([10]) was used. Textgrids were automatically created using forced alignment with an HMM-based speech recognition system. The alignments were further checked manually. A Praat script written by Xu ([11],[12]) was used to manually correct errors produced by Praat's pitch calculation algorithm. The script was also used to produce time-normalized pitch values which were used to produce Figure 1 and Figure 2.

There is no agreement about the shape of Finnish pitch accents (see [1],[2],[13]). Therefore, as an indicator of 'accentedness', we used the pitch slope defined as

$$slope = \frac{\text{2nd pitch value - 1st pitch value}}{\text{time of 2nd pitch - time of 1st pitch}} \qquad (1)$$

The pitch values were calculated in semitones, therefore (1) gives word internal slopes in semitones per second (st/s). The pitch values were measured in the middle of the first syllable nucleus and the middle of the second syllable nucleus. We did not measure peak values to avoid influences from consonant perturbations. In the case of monosyllabic verbs, the second pitch value was measured at the end of the nucleus.

Statistical analyses were done using linear mixed models as implemented in R ([14],[15]). Sentences with a verb slope differing more than two standard deviations from the mean were excluded (47 cases). In the comparisons between the verbs and the other constituents of the sentence, outliers, calculated as ±2 standard deviations around the mean, were also eliminated for the other word classes (95, 48, 39 and 54 cases for subjects, objects, pre- and postverbal adverbials, respectively).

3. Results

3.1. Pitch contours of verbs

In the data of this study, verbs with clearly falling or rising pitch, as well as those with rather flat contours occurred. The values for the slope ranged from -37.01 to 22.59 st/s. However, the median slope of -4.08 and mean slope of -4.925 indicated that most of the verbs had a falling pitch. More precisely, this is the case for 1094 of the 1505 verbs, while 411 verbs had a rising pitch. This is also reflected in the average contours in Figure 1 and Figure 2.

Figure 1 compares the time-normalized average pitch contours for verbs with different syllable structures. For the disyllabic words, the pitch difference between the first and the second syllable was considerably larger when the first syllable contained only one vowel (mean slope -6.895) than when it contained two (mean slope -2.704). The mean slope for monosyllabic verbs was -5.198, being thus closer to CV.CV-words. These results were only partly in line with previous findings for the correlation of quantity and tone on Finnish nouns ([13]). The dynamic falling pitch found on bimoraic first syllables of nouns can also be seen for the bimoraic first syllables of the verbs in Figure 1. However, the static high pitch found for monomoraic syllables in case of the nouns was not present for the verbs in our data.

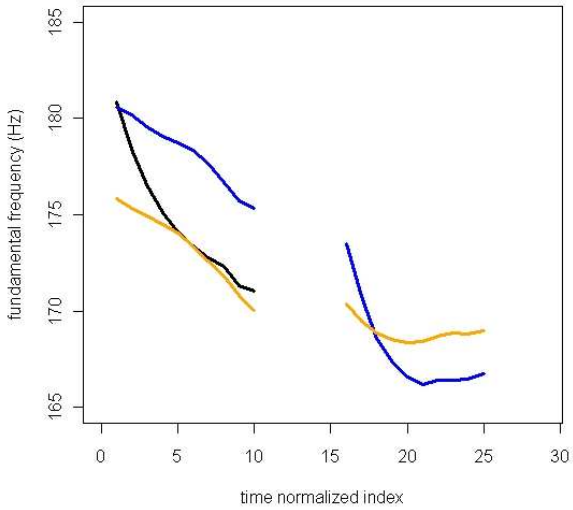

Figure 1: *Time-normalized average pitch contours of monosyllabic verbs (black line) and disyllabic verbs containing one (blue line) or two (orange line) vowels in the first syllable.*

3.2. Influences on verbal accent slope

The pitch contours of the verbs did not differ only according to their syllable structure; one further influence is illustrated in Figure 2, which shows the average pitch contours of sentences with and without adverbials. The steepest fall was found on verbs in sentences that had both a pre- and a postverbal adverbial (mean slope -7.703). The pitch fell more gradually on verbs that had either a preceding (-4.662) or a following (-3.818) adverbial. The most level pitch was observed on verbs that were directly adjacent to the subject and object (mean slope -2.462).

To assess whether this influence of the adverbial context on the pitch slope of the verbs was significant, we fitted a linear mixed effects model to the data. We also included further, phonetically and linguistically plausible predictors in the model in order to inspect their possible influence. It could for example be assumed that other factors related to the phonetic environment, like pitch range or the duration of the sentence and the separate words, have an effect. In addition, to assess whether the pitch on verbs was interpolated, we considered the influence of the pitch height of the syllables adjacent to the verb. We tested the significance of the pitch value in the middle of the last syllable of the preverbal adverbial – if present – or the subject and of the first syllable of the postverbal adverbial or the object.

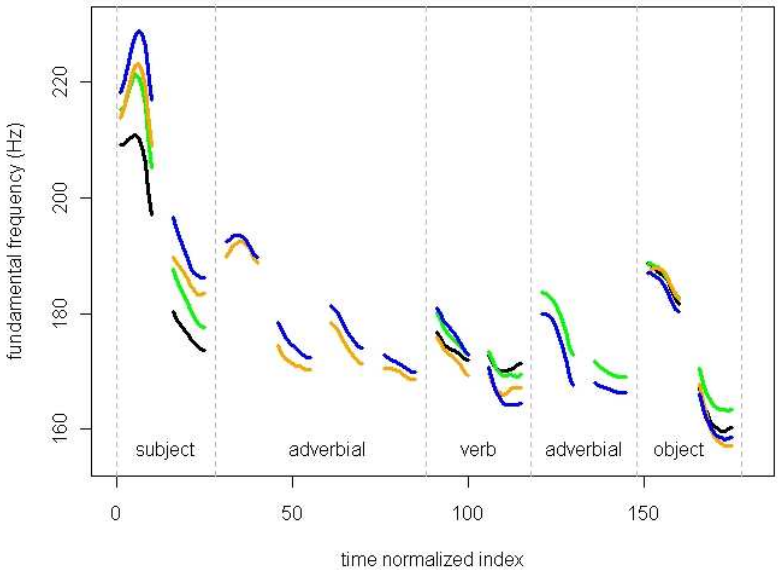

Figure 2: *Average time-normalized pitch contour of sentences with (a) an adverbial in preverbal position (orange line), (b) an adverbial in postverbal position (green line), (c) two adverbials (blue line), or (d) sentences without adverbials (black line).*

The best fit to the data was observed with a model that included

(a) sentence duration (t=2.679, p<0.01),
(b) the presence or absence of a preverbal adverbial (t=-3.881, p<0.001),
(c) the presence or absence of a postverbal adverbial (t=-3.915, p<0.001),
(d) the pitch range of the sentence (t=-4.051, p<0.001),
(e) the duration of the verb (t=4.214, p<0.001),
(f) the pitch slope of the subject (t=-6.155, p<0.001),
(g) the pitch slope of the object (t=-5.046, p<0.001),
(h) the number of syllables in the verb (t=3.415, p<0.005),
(i) the number of vowels in the first syllable of the verb (t=2.472, p<0.05).

Speaker and lexical item of the verb were included as crossed random factors ([15]).

This model only accounted for a subset of the data (978 sentences). The remaining sentences were discarded from the model because of missing values for the object pitch slope (g). These missing value were due to sentence final creaky voice. An alternative model without the factor object slope otherwise contained the same factors.

Factors included in the preliminary analyses that did not show an effect on the verb slope were the duration of subject and object and whether the adverbial in the postverbal position was one of the semantically emptier adverbs or a noun indicating a location. Also, a model that additionally included the pitch values measured on the syllables directly adjacent to the verb was more complicated and not significantly better than the present model.

3.3. Comparison between verbs and other words

The pitch contours depicted in Figure 2 show that, regardless of the presence of adverbials, the verb always had a smaller pitch range than any of the other words in the sentence. Whereas the mean slope for verbs was only -4.925, it was -5.575 for the postverbal adverbials, -9.995 for the preverbal adverbials, and even -22.192 for the subjects and -24.21 for the objects. This is depicted in Figure 3, which also shows that the standard deviation (sd) of the slope was much larger for subjects (18.221) and objects (26.284) than for verbs (8.975), with the sd for pre- and postverbal adverbials again being more similar to that of the verbs (10.653 and 11.689, respectively).

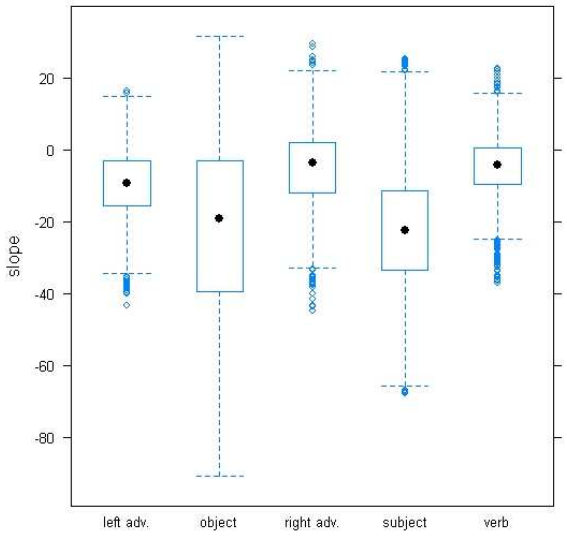

Figure 3: *Boxplot comparing the slope (in st/s) of verbs, subjects, objects and adverbials in preverbal (left adv.) and postverbal (right adv.) position.*

To test whether there was a structural difference reflected in the slope, we compared the slope of the verbs to that of the postverbal adverbs, which were most similar to the verbs in terms of accent slope. In addition, the right context adverbs were disyllabic and thus more directly comparable to the verbs (mono- or disyllabic) than the left context adverbs, which always had four syllables. A meaningful comparison also needs to consider the position in the sentence. We therefore conducted a t-test comparing the speaker means for the slope of postverbal adverbials and the speaker means for the slope of verbs in prefinal position, i.e. in sentences without right adverbial (mean slope across speakers -3.511). The t-test indicated that the difference was indeed significant (t(24)=2.793, p<0.05). This result is, however, somewhat tentative, because of differences in the syllable structure.

4. Discussion

The basic research question of this study was whether Finnish finite verbs in transitive sentences indeed do not receive accents. Specifically, we wanted to see if the pitch contours of these verbs would be shaped by interpolation.

In our data, the verbs consistently had smaller pitch excursions than the other words in the sentence. There was at least some indication that this difference would actually be particular to the grammatical function and not due to the verbs' position in the sentence. If such a difference could be confirmed, it would need a grammatical explanation such as the analysis that a West Germanic verb is integrated into the prosodic phrase of its object (cf. [8]). However, while the pitch contours of the verbs had small pitch ranges, they were neither completely flat nor determined by the pitch values of the adjacent syllables only. The answer to our research question is thus not completely straightforward.

Välimaa-Blum's formulation that unfocused finite verbs completely lack an accent ([1], likewise [2]) is the clearest testable hypothesis. It predicts that the pitch contours of verbs are completely flat, as in Välimaa-Blum's own examples, or – according to the autosegmental-metrical (AM) theory of intonation ([16],[9]) – are interpolated between the pitch targets of the neighboring words. The comparative smallness of the pitch excursions on verbs in our data could be seen as supporting this hypothesis. However, the contours on the verbs in our data were neither completely flat nor explainable by interpolation alone.

Instead, we found a consistent – if small – falling pitch slope on the majority of the verbs, which is also reflected by the average pitch contours in Figure 1. If one assumes that finite verbs do not receive accents, then the question arises of how to explain where these tonal movements come from. As they are neither likely to stem from lexical nor boundary tones, from the viewpoint of the AM-theory our data seems to suggest that the verbs do indeed carry accents.

Having said that, one also needs to account for the differences which our study found between verbs and other constituents of the sentences. Therefore, the most precise description of our data seems to be the following: Verbs are assigned accents which are systematically reduced in pitch range. Doubtlessly, this new formulation is more complicated than Välimaa-Blum's suggestion. Furthermore, it might not seem favorable to describe the structural grammatical difference between verbs and their arguments as marked by a gradual parameter such as pitch range. However, pitch range has been shown to also mark other categorical linguistic differences, such as the presence or absence of narrow focus (e.g. [11]).

We would like to point out that our analysis is somewhat tentative. To re-evaluate whether the assumption of accents is justified, we plan to investigate the phonological relevance of the pitch movements found on verbs with a perception study. Also, the assumption that the difference between verbs and adverbials is structural needs to be controlled in an experiment where syllable structure is kept identical in both categories.

5. Conclusions

This study investigated the intonation of Finnish finite verbs in transitive sentences without narrow focus. Previous descriptions of these verbs as unaccented were only partly supported. The verbs in our data did have considerably smaller pitch excursions than all the other words in the sentences, but their pitch contours were usually neither flat nor explainable by pure interpolation. Our results imply that the description that Finnish finite verbs are not assigned accents is at least problematic.

A further result, which we have not discussed in detail, was that the pitch range found on adverbials was also considerably smaller than that of subjects and objects. On the basis of our materials alone, it is impossible to say whether this was simply due to their linear position or whether the difference between arguments and adjuncts would be relevant, as has been described for some West Germanic languages (e.g. [8]).

6. Acknowledgements

We would like to thank Heini Kallio and Mona Lehtinen for their help in recording and editing the data. This research was supported by a grant from the German National Academic Foundation to the first author and two grants from the Academy of Finland (grant numbers 128204 and 107606) to the second author.

7. References

[1] Välimaa-Blum, R., "A Pitch Accent Analysis of Intonation in Finnish", Ural-Altaische Jahrbücher, 12: 82-94, 1993.

[2] Suomi, K., Toivanen, J. and Ylitalo, R., "Finnish Sound Structure. Phonetics, Phonology, Phonotactics and Prosody". Oulu University Press, 2008. Online: http://herkules.oulu.fi/isbn9789514289842/isbn9789514289842.pdf

[3] Iivonen, A., Nevalainen, T., Aulanko, R. and Kaskinen, H., "Puheen Intonaatio", Gaudeamus, 1987.

[4] Schmerling, S. F., "Aspects of English Sentence Stress", Texas University Press, 1976.

[5] Bresnan, J., "Sentence Stress and Syntactic Transformations", Language, 47:257-281, 1971.

[6] Bolinger, D., "Accent Is Predictable (If You're a Mind-Reader)", Language, 48(3): 633-644, 1972.

[7] Gussenhoven, C., "Focus, Mode and the Nucleus", Journal of Linguistics, 19:377-417, 1983.

[8] Féry, C., "German Sentence Accents and Embedded Prosodic Phrases", to appear in Dehé, N., Ishihara, S., and Feldhausen, I. [Eds], Lingua special issue on The Phonology-Syntax Interface.

[9] Ladd, D. R., "Intonational Phonology", Cambridge University Press, 1996.

[10] Boersma, P. and Weenink, D., "Praat: Doing Phonetics by Computer" (Version 5.1.09) [Computer program], 2009. Online: http://www.praat.org/.

[11] Xu, Y., "Effects of Tone and Focus on the Formation and Alignment of f0 Contours", Journal of Phonetics, 27:55-105, 1999.

[12] Xu, Y., "TimeNormalizeF0.praat" (Version 2.7.1) [Praat script], 2009. Online: http://www.phon.ucl.ac.uk/home/yi/tools.html.

[13] Vainio, M., Aalto, D., Järvikivi, J. and Suni, A., "Quantity and Tone in Finnish Lexically Stressed Syllables", in Belotel-Grenié, A. and Grenié, M. [Eds], Proceedings of the Second International Symposium on Tonal Aspects of Languages, 121-124, Université de La Rochelle, 2006.

[14] R Development Core Team, "R: A Language and Environment for Statistical Computing", R Foundation for Statistical Computing, 2009. Online: http://www.R-project.org.

[15] Baayen, R. H., "Analyzing Linguistic Data: A Practical Introduction to Statistics Using R", Cambridge University Press, 2008.

[16] Pierrehumbert, J. B., "The Phonetics and Phonology of English Intonation", Ph.D. diss., MIT, 1980.