Nordic Journal of Linguistics

Additional services for *Nordic Journal of Linguistics:*

What do compounds and noun phrases tell us about tonal targets in Finnish?

Anja Arnhold

Link to this article: http://journals.cambridge.org/abstract_S033258651500013X

How to cite this article:
Anja Arnhold (2015). What do compounds and noun phrases tell us about tonal targets in Finnish?. Nordic Journal of Linguistics, 38, pp 221-244 doi:10.1017/S033258651500013X

Request Permissions : Click here

What do compounds and noun phrases tell us about tonal targets in Finnish?

Anja Arnhold

This article compares three accounts of Finnish intonation using a perception experiment with manipulated f0 contours. The experiment involved compound/noun phrase minimal pairs differing in f0 pattern. To address the question of tonal specification, manipulations changed f0 contours of recorded compound words, associating them with f0 patterns having different components of the naturally occurring f0 rise-fall contour. Thus, the study investigated which tonal targets were crucial for the perception of a complete tonal contour inducing a noun phrase interpretation. Results suggested that the falling part of the rise-falls, modelled as realisations of a high and a following low target, was essential. They furthermore revealed evidence for these targets being associated with prosodic phrases, as well as for Finnish tonal targets being characterised by a flexibility that contrasts with accent realisations in languages like English.

Keywords accents, Finnish, intonation, perception, phrase languages, phrasing, tonal specification

Department of Linguistics, University of Alberta, Edmonton T6G 2E7, Alberta, Canada.
anja.arnhold@gmail.com

1. INTRODUCTION

Finnish intonation is generally characterised as a series of rising-falling accents, which in the default case decrease in height over the course of a sentence (see the illustration in Figure 1).[1] The relative scaling of the rise-fall contours is adjusted when a part of the sentence is focused; for example, if the sentence in Figure 1, *Moona liimaa naavaa* 'Moona is gluing lichen', answered the question 'Who is gluing lichen?', the rise-fall contour on the subject *Moona* would show a wider excursion, while the following f0 movements would be compressed (see Mixdorff et al. 2002, Vainio & Järvikivi 2007). Interestingly, however, the shape and timing of the rise-falls usually remains the same (Välimaa-Blum 1993, Mixdorff et al. 2002, Arnhold 2014). Accordingly, phonological accounts of Finnish intonation generally agree in describing accents that are, with the exception of a few marginal cases, uniformly of the same type or shape (see Välimaa-Blum 1993, Iivonen 1998, Suomi, Toivanen & Ylitalo 2008). However, different specifications of this uniform accent have been suggested.
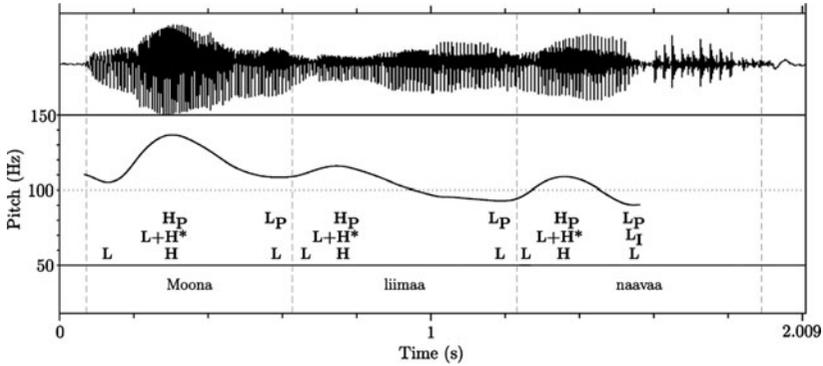
Figure 1. Simple Finnish sentence *Moona liimaa naavaa* 'Moona is gluing lichen' annotated in terms of $H_P L_P$ phrase tones (top row), L+H* accents (middle row) and LHL accents (bottom row).

Using data from a perception experiment, this article assesses the differences between the three autosegmental-metrical accounts of Finnish intonation suggested by Välimaa-Blum (1988, 1993), Suomi et al. (2008) and Arnhold (2014), illustrated in Figure 1.[2] The autosegmental-metrical framework (Bruce 1977, Pierrehumbert 1980, Ladd 1996, for a short overview see Gussenhoven 2000), characterises f0 contours as realisations of a series of high (H) or low (L) tonal targets, with the absolute f0 values of their implementation depending on factors like the individual speaker's f0 range and, crucially, the values of neighbouring targets. The framework further assumes a distinction between tones constituting accents, which are associated with prominent syllables (or, for some languages, moras), and tones associated with prosodic phrases (for example in English, the utterance *You are ill?* is marked as a question with a final pitch rise, modelled as an H tone). The first line of annotations in Figure 1 illustrates an account employing only this latter type of tones, first suggested in Arnhold (2014), and also advocated here. While autosegmental-metrical approaches differ in the number and names of the prosodic domains they assume, this article employs the phonological phrase and the higher-ranking intonation phrase. Following Hayes & Lahiri (1991), they will be referred to as P-PHRASE and I-PHRASE, respectively, in this article and subscript letters will mark the association between tones and phrases. The present account assumes that Finnish P-phrases are marked by two tones, H_P and L_P, realised early and late in the P-phrase, respectively. The f0 rise at the beginning of each word in Figure 1 is understood as a realisation of H_P and the following fall as due to L_P. In the utterance in Figure 1, each word thus forms a P-phrase of its own, but see Figures 2 and 3 below for examples of P-phrases spanning more than one word. Additionally, a final L_I tone is associated with the I-phrase, but omitted from Figure 1 and the following discussion for the sake of simplicity.

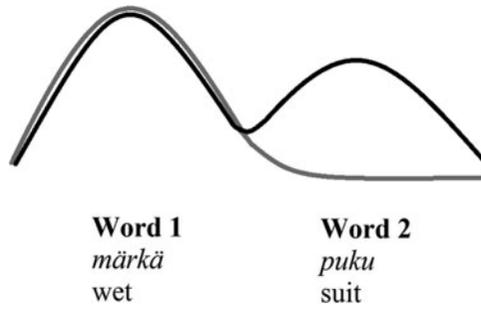| **Word 1**	**Word 2**
märkä	*puku*
wet	suit

Figure 2. Schematic illustration of f0 contours for Finnish compounds (grey line) and noun phrases (black line).

The middle annotation line in Figure 1 is based on Välimaa-Blum's (1988:104–124, 1993) account, who finds L+H* and rarely L*+H accents in her material. In autosegmental-metrical notation, the star marks a tone which is associated with the stressed syllable, while preceding ('leading') or following ('trailing') tones occur at a fixed distance from the starred tone. The f0 rises in Figure 1 are thus annotated as L+H* accents, with an H* associated with the first syllable of each word and a leading L tone realised shortly before (Finnish stress is always word-initial). The following f0 falls on the first and second word are simply transitions to the next L+H* accent under this account. For the last word, the fall is similarly understood as an interpolation between the H* target of the accent and a phrase-final L_I.

By contrast, Suomi et al. (2008:80–82) find stable alignment for both the L and the H tone as well as evidence of an L target following the H. They therefore argue that Finnish accents need to be specified as LHL, abandoning the starred notation and explicitly ruling out a description with only two tonal targets per accent, e.g. L+H*, H*+L, etc. The rise-falls in Figure 1 are accordingly fully specified by these tritonal accents, with the LH part of the accents accounting for the rises and the HL targets being realised as falls. Again, a final L_I is omitted from the annotation for simplicity; all three accounts agree in assuming that the end of an I-phrase can be marked by either L_I or H_I.

As Figure 1 illustrates, Finnish f0 contours can frequently be annotated equally well with $H_P L_P$ phrase tones, L+H* or LHL accents. On the basis of speech production data, it is difficult to motivate a preference for one of these accounts. Therefore, this article addresses the question of an optimal tonal specification for Finnish using a perception experiment involving manipulated f0 contours. Specifically, it reports on an experiment that investigated which f0 targets need to be realised on a word in order for a listener to perceive it as having a contour of its own, identifying the whole string as a noun phrase. The materials contained ambiguous items that are distinguished only by prosody as a compound or noun phrase (e.g. *märkäpuku* 'wetsuit' vs. *märkä puku*

'wet suit'). As shown by Niemi (1984:26–96), such minimal pairs are distinguished by duration, f0 and intensity in Finnish, with intensity constituting a less important cue than duration and f0. The study reported on in this article concentrated on pitch cues to the distinction. Figure 2 gives a schematic illustration of the different f0 contours.

The present account analyses the distinction between two-word compounds and noun phrases in the following way: Compounds consist of two prosodic words united in one P-phrase, while for a noun phrase, the two prosodic words each constitute a separate P-phrase. An accent-based account would state that noun phrases carry two accents, while compounds are realised with only one. Building on this distinction, the perception experiment used test stimuli with manipulated f0 contours to test which f0 movements induced the perception of a separate phrase or accent on the second word and thus of a noun phrase. In naturally occurring realisations, the second parts of the segmentally ambiguous items either do not carry a rise-fall f0 contour, marking them as compounds, or they carry a complete rise-fall, marking them as noun phrases. Again, for naturally occurring realisations, it is not possible to determine which part of this rise-fall is a realisation of tonal targets and necessary for the perception of a noun phrase and which part is just an interpolation between targets. Therefore, the present study used manipulated stimuli to test whether the f0 rise, fall, or both together are necessary for the perception of a noun phrase. Specifically, it investigated whether three tonal targets LHL, i.e. a complete rise-fall, need to be realised for native speakers to perceive a separate contour or whether 'incomplete' contours corresponding to the bitonal specifications L+H* (rising) or $H_P L_P$ (falling) have the same effect. To this end, manipulated stimuli with only the f0 rise or only the f0 fall were created in addition to contours reproducing the naturally occurring realisations.

2. METHOD

The experiment investigated the effects of pitch manipulations on speech processing with a two-alternative forced choice task. Participants were not asked which word they heard, but had to choose a picture according to an auditory instruction which included a manipulated target item (e.g. *Nyt valitse ruudulta märkäpuku/märkä puku* 'Now select the wetsuit/wet suit from the screen'). They were thus presented with a visual and an auditory stimulus in parallel.

2.1 Participants

Twenty-eight native speakers of Finnish, all university students, participated in the study. One participant, subject 20, did not complete the experiment due to technical error. The remaining 27 participants (20 female) were between 18 and 47 years
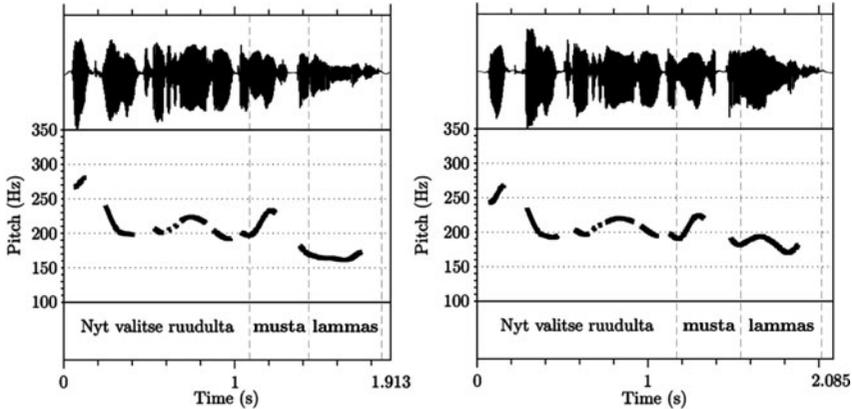
Figure 3. Waveform and pitch tracks for the original recordings of the item *musta lammas* 'black sheep' in the carrier sentence. The compound (left panel) refers to a person who ignores traditions and sticks out from a group, the noun phrase (right panel) refers to an animal.

old (mean age 25.11 years) and had all attended primary school in the Helsinki metropolitan area. One participant identified herself as bilingual with Finnish clearly being her stronger language, her other language being French. Another participant reported impaired hearing in his left ear, but this had no effect on his overall hearing ability. Since the prompts were mono sounds and input to both ears was identical, no effect on his performance was expected. Indeed, data from both of these participants showed the same patterns as for the other subjects in the group and was thus included in the evaluation. None of the other participants reported any hearing loss. All participants were compensated for their time with a shopping voucher.

2.2 Items

Target items were 34 compound/noun phrase minimal pairs of the type 'wetsuit' vs. 'wet suit'. All items had to be depictable, with the two meanings being visually clearly distinguishable (excluding pairs like *talonmies* 'janitor' vs. *talon mies* 'the man of the house'). A university student from Helsinki confirmed that she was familiar with all the items and their meanings as both compounds and noun phrases.

2.3 Auditory stimuli

A 26-year-old native speaker of Finnish from Helsinki, an advanced student of phonetics with no knowledge of the objective of the study, recorded all items in the carrier sentence *Nyt valitse ruudulta* _____ 'Now select _____ from the screen' (see Figure 3).

For each item pair, six distinct versions with manipulated f0 contours were created, as illustrated in Figure 4. The figure only shows contours of the sentence-final
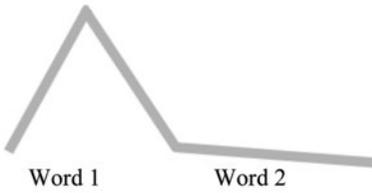
a. Compound recording

b. Noun phrase recording
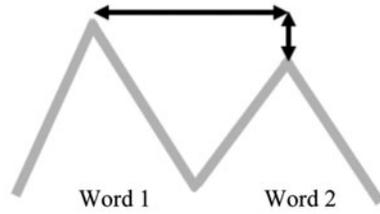
c. Manipulation 'flat'
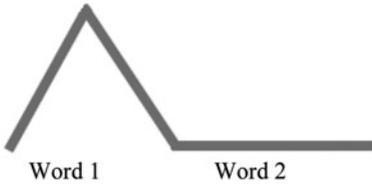
d. Manipulation 'rise-fall'
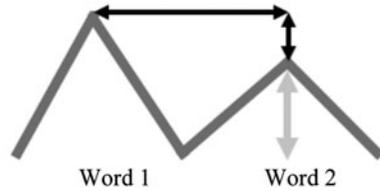
e. Manipulation 'rise'
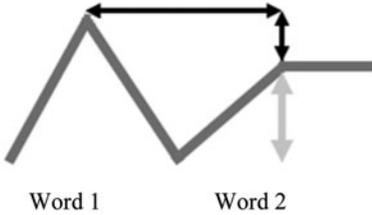
f. Manipulation 'fall'
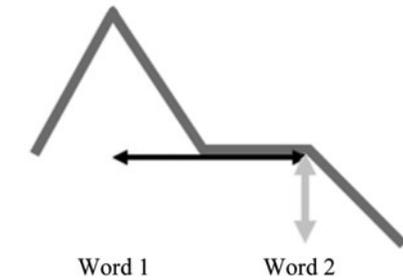
g. Manipulation 'high fall'
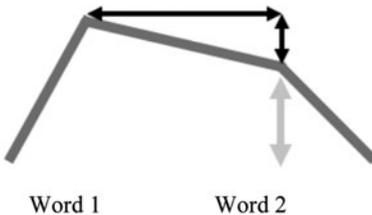
h. Manipulation 'high flat'
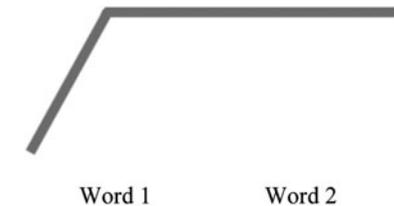
Figure 4. Schematic illustration of manipulated f0 patterns.

segmentally ambiguous compound/noun phrase, since the contour of the preceding carrier sentence was identical in all manipulations. All six manipulations were created from the original compound recording with its f0 curve stylised (Figure 4, panel a) by adding f0 movements reflecting values measured on the stylised original noun

phrase recording (Figure 4, panel b) more or less completely. Since the primary goal of the study was to address the question of tonal specification, the compound recording was chosen as the source material for all stimuli to allow a focus on the f0 cues to the perception of the compound/noun phrase distinction. All other acoustic characteristics of the compound source were retained, resulting in a very conservative design. In other words, as all other acoustic cues pointed to a compound interpretation, all occurring perceptions of these manipulated stimuli as noun phrases could truly be ascribed to the f0 cues under investigation.

The manipulations were designed as more or less complete recreations of the noun phrase contour. For the manipulation 'flat' (Figure 4, panel c), the first word of the ambiguous stretch, e.g. *märkä* 'wet' in *märkäpuku* 'wetsuit', was left unaltered, while the second word, e.g. *puku* 'suit' in *märkäpuku* 'wetsuit', carried completely level pitch. This contour was intended as a reproduction of natural compound realisations, carrying no part of the rise-fall characterising noun phrase realisations. For the manipulation 'rise-fall' (Figure 4, panel d), both the rise and the fall of a typical noun phrase realisation were present on the second word. The f0 peak that was added to create this movement had the same temporal and f0 difference from the peak on the first word as in the original noun phrase recording, as symbolised by the black arrows in Figure 4. In terms of segmental alignment, the peaks were realised on the first syllable vowel of the second word (i.e. the first *u* of *puku* in *märkäpuku*) for all target items, while the falls always ended at the end of the word.

In the manipulation 'rise' (Figure 4, panel e), the second word carried a pitch rise with the same timing and magnitude as for the 'rise-fall' contour, but f0 remained level instead of falling after the peak. Thus, it contained only the rising part of the rise-fall typically realised on the second parts of noun phrases, while the following f0 fall was missing. Conversely, the manipulation 'fall' (Figure 4, panel f) contained only a fall corresponding to that of the 'rise-fall' contour in timing and magnitude (indicated by a grey arrow). However, it was preceded by level f0 and thus reached a much lower end point. Therefore, an f0 fall of the same magnitude and timing and on the same level as the one in the 'rise-fall' version appeared in the manipulation 'high fall' (Figure 4, panel g). Instead of a rise, this fall was preceded by a high, slightly tilted plateau. Thus, the contour of the first word, e.g. *märkä* 'wet' in *märkäpuku* 'wetsuit' was altered as well, in contrast to the first four manipulations. Finally, in the manipulation 'high flat' (Figure 4, panel h), f0 was completely level following the rise on the first word. All manipulated versions were resynthesised using PSOLA (Pitch Synchronous Overlap and Add) implemented in Praat (Boersma & Weenink 2010).

2.4 Fillers

In addition to these target items, which were segmentally ambiguous between noun phrases and compounds, 34 filler item pairs were selected. These consisted of

an unambiguous noun phrase and an unambiguous compound. To make picture selection non-trivial even for the filler trials, the members of the filler item pairs were phonetically similar, as in *mustekala* 'squid' vs. *musta kala* 'black fish', or visually very similar, as for the pair *raitiovaunu* 'tram car' vs. *vihreä juna* 'green train'. Furthermore, to make item and filler sets more similar, each filler item pair was matched to a target item pair with a similar compound lemma frequency per million, calculated on the basis of the Karjalainen lexical database.[3] Filler compounds were also overall roughly matched to the targets in terms of word length.

For the auditory stimuli, fillers were recorded embedded in the same carrier sentence as the target items. They were resynthesised as close approximations of the original pitch contours. In contrast to the target items, auditory stimuli unambiguously identified either the compound or the noun phrase referent as the correct choice, inducing both choices equally often. Pictures of the filler items were shown six times per session, so that participants would encounter them exactly as often as the pictures of the target item referents.

2.5 Procedure

Stimulus presentation and data collection used E-prime 1.1 and an E-Prime Serial Response Box, collecting response time and choice. Participants were told that two pictures would appear on the screen and that their task would be to select the picture on the left or on the right by pressing the corresponding button according to instructions from headphones. They were asked to respond as fast, but as accurately, as possible.

For each trial, the participant first saw a fixation cross displayed in the middle of the screen for 1000 ms. Subsequently, the stimulus slide appeared, showing the two pictures representing the noun phrase and the compound referent of the item pair next to each other, surrounded by a white frame. All pictures were landscape format photographs of the same size. For some pairs, an arrow was used in both pictures to point out a particular detail for clarification. To avoid priming, the location of the pictures was balanced so that the photograph showing the compound referent was always displayed on the left for half of the items and on the right for the other half. The auditory stimulus was presented at the same time over Sennheiser HD 515 headphones. The onset of the sound file was time-aligned with the appearance of the visual stimuli. Thus, the pictures were visible to the participant between 1338 ms and 1609 ms before the onset of the target word, depending on the speech rate of the individual frame sentence. Response was enabled from the appearance of the stimulus slide. The pictures remained visible after the offset of the auditory stimulus, staying on the screen until the subject responded, but maximally for 4000 ms altogether.

Each experimental session started with six filler-type practice trials that did not occur again later in the session. The experimental session proper consisted of two blocks, separated by a break: The first block contained all experimental items in the

	Tonal targets		
Predictions	H_PL_P	L+H*/L*+H	LHL
Higher proportion of noun phrase responses	Rise-fall High fall	Rise-fall Rise	Rise-fall
Lower proportion of noun phrase responses	Rise Fall High flat Flat	Fall High fall High flat Flat	Rise Fall High fall High flat Flat

Table 1. Summary of predictions based on three accounts of Finnish intonation.

contour conditions 'flat', 'rise-fall', 'rise' and 'fall' (see Figure 4), interspersed with two repetitions of the filler items in the noun phrase version and two repetitions of the fillers indicating the compound as the correct choice (altogether 272 trials). The second block included all experimental items in the pitch conditions 'high flat' and 'high fall' together with another two repetitions of the fillers, once with the noun phrase and once with the compound as the correct choice (136 trials). The intention in splitting the materials up in this way was to separate the possibly confusing patterns with manipulations affecting the first part of the compound/noun phrase from the rest of the materials. Within each block, trial order was pseudo-randomised, creating four lists.

3. HYPOTHESES

As described in the introduction, Finnish two-word compounds are assumed to be realised as one P-phrase spanned by the realisation of H_PL_P tones, while two-word noun phrases are assumed to be regularly realised as two P-phrases with separate contours on both words – or, in accent-based approaches, compounds are analysed as bearing one initial accent, while in noun phrases each of the two words bears an accent of its own (see Figures 2 and 3 above). The hypothesis for the present study is that a realisation of a segmentally ambiguous compound/noun phrase like *märkäpuku* 'wetsuit' will more frequently be perceived as a noun phrase if all tonal targets associated with a P-phrase – or accent – are realised on the second word (e.g. *puku* 'suit'). Given this hypothesis, the different accounts of Finnish intonation generate different predictions, as summarised in Table 1, since they assume different tonal targets.

Figure 5 displays the six manipulated f0 contours used in the stimuli (see Figure 4) with three different annotations. The first row shows renderings of the
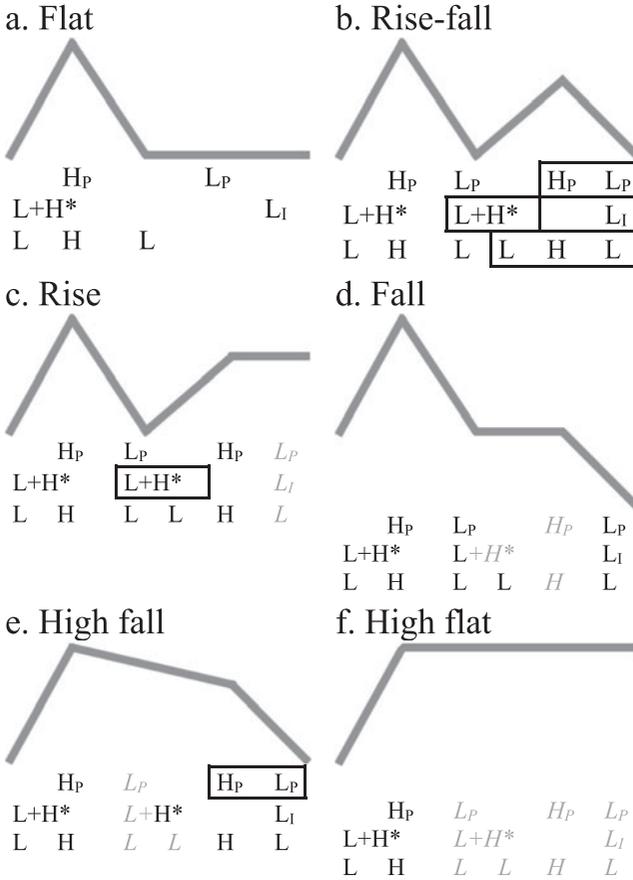
a. Flat

Hₚ	Lₚ	
L+H*		Lᵢ
L H	L	

b. Rise-fall

		Hₚ Lₚ
Hₚ	Lₚ	L+H*
L+H*	L+H*	Lᵢ
L H	L	L H L

c. Rise

Hₚ	Lₚ	Hₚ	*Lₚ*
L+H*	L+H*		*Lᵢ*
L H	L L	H	*L*

d. Fall

Hₚ	Lₚ	*Hₚ*	Lₚ
L+H*	L+*H**		Lᵢ
L H	L L	*H*	L

e. High fall

		Hₚ Lₚ
Hₚ	*Lₚ*	Lᵢ
L+H*	*L+H**	
L H	*L L*	H L

f. High flat

Hₚ	*Lₚ*	*Hₚ*	*Lₚ*
L+H*	*L+H**		*Lᵢ*
L H	*L L*	*H*	*L*

Figure 5. Manipulated f0 contours and tonal targets according to an annotation in terms of HₚLₚ phrase tones (top row), L+H* accents (middle row) and LHL accents (bottom row). Tonal targets not realised in the f0 contour are shown in grey italics, complete realisations of tonal targets associated with the second word are boxed.

tonal targets suggested in the present account, i.e. HₚLₚ tones associated with the P-phrase. The second row employs Välimaa-Blum's (1993) L+H* accents, while the third gives an annotation in terms of LHL accents, following Suomi et al. (2008). These annotations mark tones corresponding to minima or maxima in the contours in black, while missing or unrealised tones are marked in grey italics. Black boxes highlight annotations of complete tonal contours on the second word, i.e. contours that realise all tonal targets according to the respective analysis. These conditions are expected to induce significantly more noun phrase responses.

For the two contours resembling naturally occurring two-word compounds ('flat', panel a) and noun phrases ('rise-fall', Figure 5, panel b), all three accounts generate

the same predictions. The 'flat' contour is clearly expected to lead to a relatively low proportion of noun phrase choices, whereas the 'rise-fall' contour should lead to a significantly higher percentage of noun phrase responses. Also, all three annotations agree in marking the second word as not having a complete tonal contour of its own for the 'high flat' contour and the 'fall' contour, thus predicting a low proportion of noun phrase choices induced by these manipulations. For 'rise' and 'high fall' contours, the predictions of the three analyses differ.

The analysis in terms of $H_P L_P$ tones interprets the 'rise-fall' contour on the second word as constituting a P-phrase on its own and thus predicted a relatively frequent noun phrase perception for this contour. In contrast, the 'rise' contour does not constitute a complete second P-phrase associated with $H_P L_P$ tones in this view, since a realisation of the crucial L_P tone is missing. Similarly, the 'fall' contour is analysed as incomplete. Here, the rise-fall contour spanning the preceding P-phrase (the first word of the compound/noun phrase) is followed by a low plateau and a pitch fall signalling the presence of another L_P tone, but there is no preceding rise and thus no indication of an H_P realisation.[4] Both the 'rise' and the 'fall' manipulation were thus expected to induce relatively few noun phrase responses. The 'high fall' contour, in contrast, is understood as marking the presence of both H_P and L_P on the target word, while the realisation of the $H_P L_P$ tones of the preceding P-phrase is defective, resulting in the lack of an intervening pitch fall. Therefore, listeners were expected to interpret this contour as signalling a noun phrase interpretation significantly more often than the 'flat' pattern. Likewise, the contour of the first part of the compound/noun phrase is realised incompletely for the 'high flat' contour. Since this contour gives no indication of a second P-phrase, the interpretation as a compound should be unambiguous.

Following the claim by Suomi et al. (2008) that Finnish accents need to be specified as a tritonal LHL instead of just positing targets for either the rise (LH) or the fall (HL), one would assume that only the 'rise-fall' manipulation should induce a significant increase in noun phrase responses. The 'rise', 'fall' and 'high fall' contours should induce a relatively low proportion of noun phrase choices, although the presence of incomplete accents could also induce confusion and a certain randomness of responses. The 'high flat' contour signals the presence of one incompletely realised accent on the first part according to this approach and should thus clearly result in a low number of noun phrase choices.

Following Välimaa-Blum's (1993) description of accents as L+H* (or marginally L*+H), the 'rise-fall' contour should also be the one clearly eliciting noun phrase responses, since participants would expect the accentual rise to be combined with a fall to a phrase-final L_I boundary tone. However, the accent should still be perceivable without the boundary tone so that the 'rise' pattern should lead to the same proportion of noun phrase responses as does the 'rise-fall' manipulation. In contrast, the 'fall' and 'high fall' manipulations, lacking the rising pitch movement characteristic of

Response	Flat	Rise-fall	Rise	Fall	High fall	High flat
Compound	632 (71%)	530 (59%)	567 (63%)	624 (70%)	485 (54%)	663 (74%)
NP	258 (29%)	365 (41%)	332 (37%)	265 (30%)	413 (46%)	229 (26%)

Table 2. **Number and percentage of responses selecting compound vs. noun phrase (NP) referents for the different f0 contour conditions.**

an accent, should induce significantly fewer noun phrase responses according to Välimaa-Blum's account of Finnish intonation.

Finally, a predominance of compound responses was anticipated across the experimental conditions, irrespective of the intonational account. The auditory stimuli were based on compound recordings. Segmental timing, as well as all other phonetic cues to the compound interpretation – apart from f0 – were retained, so that the materials contained a strong bias towards the compound choice. This set a high bar for the assumption that by manipulating the pitch contour alone, it is possible to create the perception of a noun phrase on the basis of a compound recording. As mentioned above, the experiment was purposefully set up to be conservative in this respect, since the aim of the study was to compare different analyses of f0 patterns, not to test the cues to the perception of the compound/noun phrase distinction per se (see Niemi 1984 for more on the latter topic). Accordingly, the predictions spelled out above concern relative differences expected between conditions, not absolute differences as regards the number of noun phrase responses.

4. RESULTS

Of the overall 5508 experimental trials conducted (34 lexical items × 6 f0 contour conditions × 27 participants), 96 data points (2%) were lost due to participants failing to respond before the presentation automatically proceeded to the next trial. Additionally, responses earlier than 200 ms after the target onset were removed, eliminating 49 trials (1% of the collected data). Thus, 5363 responses were evaluated.

The participants' choices of the pictures showing the referent of the compound or of the noun phrase interpretation of the ambiguous sequences appear in Table 2. As expected due to the compound bias inherent in the materials, subjects chose the compound interpretation in the majority of trials across the different conditions (3501 cases, i.e. 65% of the responses overall). In spite of this bias, there was also a sizeable proportion of noun phrase responses. Crucially, the distribution of responses varied between the experimental conditions. While the 'high fall' condition elicited noun phrase responses in almost half of the cases, the two flat conditions led to participants choosing the noun phrase referents about 20% less often.

To test the significance of these differences, the data were analysed by fitting linear mixed-effects models as implemented in the statistical analysis software R (Baayen, Davidson & Bates 2008, R Development Core Team 2011), using the 'lmer' function in the package 'lme4' (Bates, Maechler & Bolker 2011). Since the choice between noun phrase and compound interpretation was a binary dependent variable, these models were binomial and the function automatically calculated *p*-values on the basis of the *z*-score. Linear mixed-effects models are a generalisation of ordinary logit models (or logistic regression) and inherit many of their advantages, especially in the analysis of categorical data: They can include categorical as well as numerical predictors, provide information on size and directionality of effects and perform well with unbalanced data sets (Jaeger 2008). In contrast to ordinary logit models, they allow the specification of random effects to take into account the connection between data points coming from the same subject or item in a repeated measures design like the present one.

As in normal logistic regression, linear mixed-models estimate the (log transformed) odds of outcomes – in this case the participants' choice of compound vs. noun phrase responses – with a linear function that contains predictor variables plus a random error term and (optionally) an intercept (see Baayen 2008:214–216; Jaeger 2008 on the advantages of the log-odds transformation). Here, the intercept was retained as is the default for the lmer function. Models including different predictor variables and random effect structures were fit and the model with the best fit to the data was determined with an ANOVA comparison of log-likelihood. The crucial predictor variable was the experimentally manipulated f0 contour on the segmentally ambiguous compound/noun phrase. In addition, the following predictor variables were tested: logarithmically transformed response time, logarithmically transformed compound lemma frequency per million, logarithmically transformed lemma frequency of the first word per million and logarithmically transformed lemma frequency of the second word per million (calculated on the basis of the Karjalainen database, see endnote 3). Tested random factors were subject, item and trial. Only variables significantly improving model fit were retained. The best-fitting model for the entire data set, containing f0 condition and reaction time as predictors and subject and item as random effects, is shown in Table 3. Additionally, this process was performed separately for the responses to the manipulations with non-flat f0 on the second word, i.e. 'rise-fall', 'rise', 'fall' and 'high fall'. The best model for this subset, with the same predictors and random effects, appears in Table 4. The table captions give the number of observations as well as the log-likelihood, describing the goodness of fit of the model (for more information see Baayen 2008, Jaeger 2008).

For the model of the whole data set, the 'flat' condition with a low, completely flat pitch contour on the second word of the ambiguous stretch was the intercept, since it was designed to approximate the f0 contour of naturally occurring compounds. The first value in the second column of Table 3 shows the log-odds of a noun phrase

Predictor	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	−5.01654	0.73355	−6.839	7.99e-12	***
Contour 'Rise fall'	0.91862	0.33623	2.732	0.00629	**
Contour 'Rise'	0.63139	0.30367	2.079	0.03760	*
Contour 'Fall'	0.17651	0.19703	0.896	0.37034	
Contour 'High fall'	1.27265	0.42937	2.964	0.00304	**
Contour 'High flat'	−0.13750	0.18317	−0.751	0.45287	
Log response time	0.52782	0.09483	5.566	2.61e-08	***

Significance codes: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Table 3. **Fixed effects for best-fitting model of responses in the complete data set (N = 5363; log-likelihood = −2510).**

Predictor	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	−4.9425	0.8269	−5.977	2.27e-09	***
Contour 'Rise'	−0.2897	0.1462	−1.982	0.04751	*
Contour 'Fall'	−0.7318	0.2478	−2.953	0.00315	**
Contour 'High fall'	0.3244	0.2080	1.560	0.11888	
Log response time	0.6554	0.1134	5.780	7.47e-09	***

Significance codes: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Table 4. **Fixed effects for best-fitting model of responses to non-flat manipulated contours (N = 3581; log-likelihood = −1747).**

response (as opposed to a compound response) for this condition as estimated by the model. Since the log-odds are the natural logarithm of the odds, the estimated odds of a noun phrase response in the 'flat' condition can be calculated as $e^{-5.01} \approx$ 0.006 and the probability as 0.006/(1+0.006) ≈ 0.006. The large negative estimate thus corresponds to very low odds, in line with the low number of noun phrase responses observed in the 'flat' condition (Table 2 above; note that this probability does not directly equal the observed proportion of noun phrase responses, since the model estimates the odds while taking the effect of reaction time and by-speaker and by-item variation into account). The other values in this column give the estimated difference of the odds of a noun phrase response in the other conditions. Positive estimates indicate an increase in noun phrase responses relative to the intercept. Thus, the estimates provide a direct indication of effect size and directionality, while the standard errors give an idea of the variation in the data and are used in the calculation of the *p*-values (Jaeger 2008).

In contrast to the 'flat' condition, participants chose noun phrase referents more often for almost all other pitch contours. This can be seen both from the percentages in Table 2 and from the positive estimates in Table 3. As expected, this effect was significant for the 'rise-fall' contour, which was designed to correspond to natural

noun phrase realisations. The increase in noun phrase choices when only an f0 rise was present on the target ('rise' contour) was likewise significant compared to the 'flat' contour. However, the 'rise' manipulation triggered significantly less noun phrase responses than the 'rise-fall' contour. This emerges from the model in Table 4 comparing the non-flat pitch contours with the 'rise-fall' as the intercept. It indicates a significant negative effect for the 'rise' manipulation. Note that the *p*-values associated with the effects of the 'rise' contour are close to the threshold of .05 both in Table 3 and in Table 4, suggesting that this manipulation neither patterned with the 'flat' nor with the 'rise-fall' condition, but fell in between the two.

A fall from the low pitch level reached at the end of the preceding word alone ('fall' pattern) did not induce a significantly larger proportion of noun phrase choices when compared to the 'flat' pattern (see Table 3) and resulted in significantly less noun phrase choices than the 'rise-fall' contour (see Table 4). That is, the effect of this manipulation on the participants' responses was more comparable to that of the 'flat' contour, leading to a clear majority of compound responses (see Table 2 again).

The f0 contour leading to the highest number of noun phrase choices was the 'high fall' condition, where the fall occurred from a high plateau sustained after the peak of the preceding word. The effect of this fall from a high level was significant compared to the 'flat' condition, and it did not differ significantly from the 'rise-fall' pattern (see Tables 3 and 4, respectively). By contrast, the 'high flat' manipulation, containing only a high plateau stretching from the f0 maximum of the first word of the segmentally ambiguous item to the end of the sentence, lead to a slightly lower number of noun phrase responses than the 'flat' contour, although the difference was not significant (see Tables 2 and 3).

Additionally, both models indicated that noun phrase choices were significantly more frequent in trials with longer response times (see the positive estimates in the model summaries). Lemma frequencies of the compound, first word and second word did not significantly improve the models and were thus not included. While there was thus no indication that lexical frequencies overall systematically affected the choice of a compound vs. noun phrase referent, the models took variation between the lexical items into account by including random by-item effects, shown in Tables 5 and 6. The specification of this random effects term significantly improved the fit to the data for both models, as did the inclusion of by-subject random effects. The by-subject random effects were further improved by including the experimental condition in the term. This indicates that participants not only differed in their general predisposition towards compound vs. noun phrase interpretations, but that they additionally differed in their response to the different manipulated f0 contours.

Figure 6 illustrates this by-subject variation in responses to the experimental conditions. It shows that while most participants selected the compound response for the majority of the trials, this tendency was more pronounced for some participants (e.g. subjects 1, 18 and 25) than for others (e.g. 7, 17 and 27). Over and above this

Random effect	Variance	Std. deviation	Correlation with random effect for:				
			Intercept	'Rise fall'	'Rise'	'Fall'	'High fall'
Item intercept	1.97503	1.40536					
Subject intercept	0.91103	0.95448					
Subject 'Rise fall'	2.61674	1.61763	−0.761				
Subject 'Rise'	2.06078	1.43554	−0.668	0.975			
Subject 'Fall'	0.61048	0.78133	−0.596	0.777	0.673		
Subject 'High fall'	4.51553	2.12498	−0.765	0.913	0.803	0.908	
Subject 'High flat'	0.43646	0.66065	−0.260	0.168	0.181	0.468	0.153

Table 5. **Random subject and item effects and correlations in the best-fitting model of responses in the complete data set.**

Random effect	Variance	Std. deviation	Correlation with random effect for:		
			Intercept	'Rise'	'Fall'
Item intercept	1.85746	1.3629			
Subject intercept	1.12256	1.0595			
Subject 'Rise'	0.18301	0.4278	−0.206		
Subject 'Fall'	1.24953	1.1178	−0.753	0.354	
Subject 'High fall'	0.76160	0.8727	0.106	−0.852	0.073

Table 6. **Random subject and item effects and correlations in the best-fitting model of responses to non-flat manipulated contours.**

general preference, the participants differed in their sensitivity to the experimental manipulations, with some of them showing a large variation across the conditions (e.g. 2, 7 and 8) and others displaying a stable proportion of compound responses (e.g. 12, 25 and 26). Notice also that three subjects (5, 13 and 24) chose the compound referent for about 50% of the trials in all conditions, suggesting that these participants responded randomly. To assess whether this introduced noise affecting the results, the models in Tables 3 and 4 were re-fit to subsets of the data excluding responses from these three participants. However, the resulting models showed the same effects to be significant.

The models summarised in Tables 3 and 4 already take by-subject and by-item differences in responses into account. As mentioned above, this is achieved through
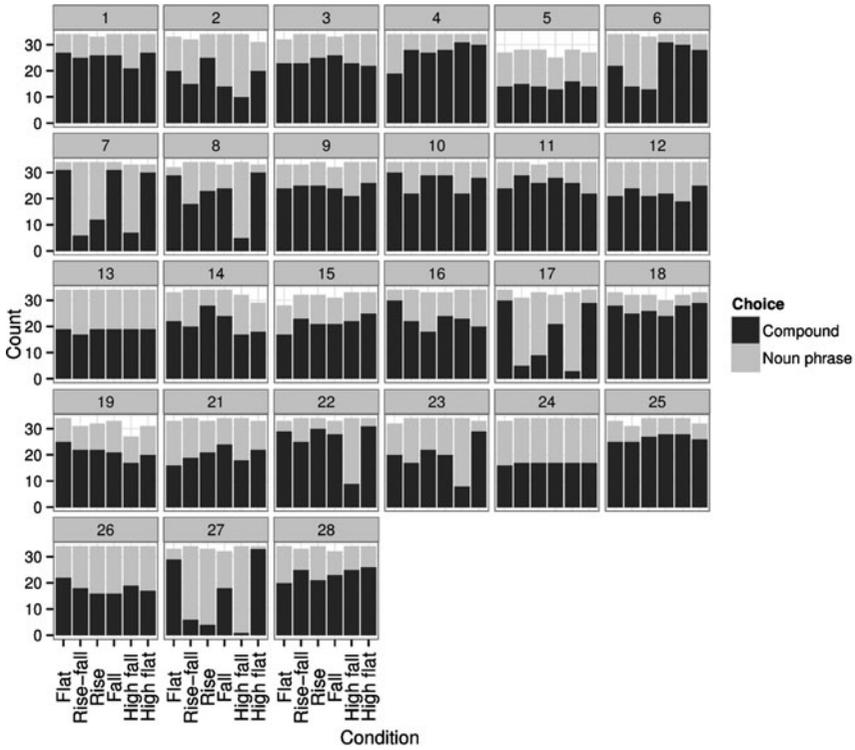
Figure 6. Distribution of compound and noun phrase choices for the different manipulated f0 contour conditions by subject.

the specification of random effects, shown in Tables 5 and 6. Random effects are defined as normally distributed around a mean of zero, with the variance of their distribution estimated by the model. The summaries in Tables 5 and 6 therefore only give the variance, standard deviations and correlations of the random effects, since the individual adjustments are not formally parameters of the model (see Baayen et al. 2008 for more details on random effects in linear mixed-effects models). Otherwise, the interpretation for the random effects summaries is similar to the interpretation of fixed effects. Since the data is modelled with a linear function, the values have a straightforward geometrical interpretation in log-odds space. Accordingly, random effects are divided into adjustments of intercepts and slopes. As stated above, the model for the complete data set as well as the model of responses to the non-flat contours had the same random effects structure. Both fitted random intercepts for subjects as well as items, and additionally specified random by-subject slopes for the f0 manipulation conditions. Thereby, the intercepts model general differences between subjects or items. Thus, the by-item random intercepts reflect the fact that some items generally elicited more noun phrase responses, while the by-subject

intercepts accounted for the fact that some subjects were generally more prone to noun phrase responses than others, as seen in Figure 6. Both tables indicate that variance in by-item intercepts was larger than variance in intercepts representing by-subject variation, which was additionally modelled with random slopes. The slopes accounted for the differences in how different subjects reacted to the experimental conditions. Table 5 suggests that compared to the 'flat' condition, subjects varied much more in how they responded to the 'rise-fall', 'rise' and especially to the 'high fall' condition, as shown by the larger values for variance and standard deviation for the random slopes associated with these conditions. For the comparison of the other non-flat contours with the 'rise-fall' pattern corresponding to natural noun phrase realisations, Table 6 displays smaller differences between the conditions in terms of the variation between subjects, although including random slopes also improved the fit for this model significantly. Finally, the right part of both tables gives a correlation matrix for the different f0 manipulation conditions. For the model of the complete data set in Table 5, all non-flat conditions showed strong negative correlations with the intercept adjustments. This reflects the fact that subjects who showed a very strong overall bias towards compound responses frequently were not much affected by the difference between f0 conditions (e.g. subjects 1, 3, 9), while subjects with larger differences between the conditions had an overall lower percentage of compound responses. Additionally, there were large positive correlations between all the non-flat conditions, indicating for example that subjects who gave a higher number of noun phrase responses for the 'rise-fall' condition also chose the noun phrase interpretation more frequently for the 'high fall' condition. Comparing only the non-flat conditions to each other, there was only one strong correlation among the slopes (Table 6): Variance in responses to the 'rise' and the 'high fall' conditions were negatively correlated, probably stemming from the fact that a few subjects gave a relatively high number of noun phrase responses to the 'high fall' condition and gave the noun phrase response relatively infrequently for the 'rise' contour (e.g. subjects 2, 8, 22). However, the random effects mostly reflected a difference in the subjects' overall responsiveness to the experimental manipulations.

Generally, the ability to take by-subject and by-item variation into account is one of the greatest advantages of linear mixed-effects modelling. Thus, the significant fixed effects appearing in Tables 3 and 4 can be said to be significant overall, i.e. after the variation between subjects and items is taken into account (Jaeger 2008:444).

5. DISCUSSION

This experiment investigated whether different manipulated f0 contours could elicit a perception of the second word of a segmentally ambiguous compound/noun phrase as carrying a complete tonal contour of its own and thus a categorisation of the

Predictions	Tonal targets			Results
	H_PL_P	L+H*/L*+H	LHL	
Higher proportion of noun phrase responses	Rise-fall	Rise-fall	Rise-fall	High fall
	High fall	Rise		Rise-fall
Lower proportion of noun phrase responses			Rise	Rise
	Rise	Fall	Fall	Fall
	Fall	High fall	High fall	Flat
	High flat	High flat	High flat	High flat
	Flat	Flat	Flat	

Table 7. Summary of results regarding choice of compound or noun phrase interpretations (right) compared to the predictions based on three accounts of Finnish intonation (left, compare Table 1).

complete string as a noun phrase. In spite of an overall bias towards compound responses, the experimental design was successful in assessing differences between the six contours (see Figure 5 above) – the patterns 'flat', 'rise-fall', 'rise', 'fall', 'high fall' and 'high flat'. Compared to the flat f0 naturally occurring on the second word of compounds, the 'rise-fall' pitch pattern naturally occurring for noun phrases, as well as the 'high fall' pattern elicited significantly more noun phrase responses (see the overview in Table 7). By contrast, the 'high flat' and 'fall' contours did not differ significantly from the 'flat' pattern, but resulted in significantly fewer noun phrase responses than the 'rise-fall' contour. The 'rise' contour appears between these two groups in Table 7, as it did trigger significantly more noun phrase choices than the 'flat' contour, but also significantly fewer than did the 'rise-fall' contour.

Table 7 further compares these results to the predictions from three accounts of Finnish intonation: first, the analysis in terms of H_PL_P phrase tones advocated here, second, Välimaa-Blum's (1988, 1993) inventory containing L+H* and L*+H accents and third, Suomi et al.'s (2008) description of LHL accents. As detailed in Section 3 above, each of the three approaches led to the prediction that some of the pitch patterns would induce more noun phrase responses and others less. The contrast between more noun phrase choices for the 'rise-fall' pattern and less noun phrase choices for target words with completely flat pitch was expected on the basis of all three accounts. With respect to the task at hand, it did not make a difference whether the pitch contour of the relevant word was flat on a low or on a high level. The analysis in terms of H_PL_P phrase tones hypothesised that only the 'rise-fall' and the 'high fall' pattern should lead to substantially more noun phrase choices, while the 'rise' and 'fall' contours should cluster with the 'flat' and 'high flat' pattern in inducing mostly compound responses. These hypotheses were largely in agreement with the results, as Table 7 illustrates: The 'rise-fall' and 'high fall' pitch patterns did indeed elicit significantly more noun phrase responses than the 'flat' contour, while the 'fall' and

'high flat' manipulations did not. Only the 'rise' contour did not fit this predicted binary classification quite as neatly as the other manipulations.

The analysis in terms of rising L+H* or L*+H accents predicted that only the 'rise-fall' and the 'rise' pattern should increase the proportion of noun phrase choices compared to the 'flat' contour, while the 'fall', 'high fall' and 'high flat' pattern should not. The results provided some evidence contradicting these predictions. Crucially, the 'rise' pattern did not induce a clear increase in noun phrase categorisations. This finding does not fit the assumption that the rising part of Finnish rising-falling pitch contours is a tonally specified accent, whereas the fall is just a transition to the leading tone of the next accent or boundary tone, as suggested by Välimaa-Blum (1993). Moreover, the high proportion of noun phrase responses induced by the 'high fall' pattern – slightly exceeding that caused by the 'rise-fall' contour – is difficult to reconcile with this approach. The 'high fall' manipulation only realised a pitch fall from a high level on the second word of the ambiguous compound/noun phrase stretch. This fall was preceded by high plateau and, before that, a rise at the beginning of the first word of the compound/noun phrase. In terms of L+H* accents, the first part would thus carry a completely realised accent, while the second part would not, as annotated in Figure 5 above. Thus, this manipulation should have led to compound identifications just as unambiguously as the 'flat' contour, although possibly some uncertainty might have been induced by the high plateau. Instead, this contour, lacking the – under the L+H*/L*+H analysis crucial – pitch rise on the last word, induced the highest proportion of noun phrase categorisations.

According to Suomi et al.'s (2008) assertion that Finnish accents are uniformly of the type LHL, it was predicted that only the 'rise-fall' pattern would lead to a higher proportion of noun phrase responses. This hypothesis was not born out. Instead, one would be lead to conclude that even incomplete accent realisations were sufficient for participants to detect the presence of an accent and accordingly choose the picture of the noun phrase referent in the present study. The patterns 'rise' and 'fall', which would be seen as 'incomplete accents', induced fewer noun phrase choices than the 'complete rise-fall'. Crucially, the 'high fall' contour also lacking the pitch rise did not differ significantly from the 'rise-fall' contour and triggered (insignificantly) more noun phrase choices than the supposedly complete accent realisation.

In sum, regarding the question of tonal specification, the analysis in terms of $H_P L_P$ seemed to best account for the data, since the pitch patterns eliciting the clearest increase of noun phrase responses were those realising a high and following low tone on the final word, whereas the presence or absence of a preceding low tone was not crucial.

In addition to bearing on the question of rising vs. falling tonal specification, the present data also support the analysis of the tonal targets as associated with prosodic phrases instead of constituting accents. Crucially, the study's findings suggest a flexibility of tonal targets. In this vein, a major observation is that in spite of the

strong compound bias in the stimuli, participants frequently chose noun phrase interpretations in response to 'incomplete rise-fall' contours. Even for the 'rise' and 'fall' contours, the number of noun phrase categorisations was not very high, but at 332 and 265 cases not exactly low, either.

Participants' reactions to a variety of incomplete tonal contours indicate a difference from the perception of accent in intonation languages. For example, speakers of a Germanic language generally either perceive an accent being realised on a certain word or they do not perceive this accent being present on this word – this kind of categorical distinction has for example been exploited in perception experiments rather similar to the current one, which have been used as evidence pertaining to the on-ramp vs. off-ramp debate (see e.g. Gussenhoven 2008, Chen 2011). In contrast, the participants of the present study were more likely to categorise the ambiguous stretch as a noun phrase when any kind of pitch movement was present on the second prosodic word. This kind of variability is more in line with a phrasing account as suggested here than with an analysis in terms of accents. Whereas no all-or-none effect in line with an accent interpretation could be found, the results are consistent with the following interpretation: When participants perceived more pitch movements than expected for the realisation of one P-phrase, they were more likely to attribute the superfluous movements to the existence of a second P-phrase and thus to choose the noun phrase interpretation.

Moreover, it is worth noting that most of the manipulations resulted in contours that do not feature in previous descriptions in the literature and might not naturally occur in Finnish. However, several native speakers of Finnish reported that none of the pitch patterns used in this experiment sounded weird or distinctly unnatural and that even ranking their (un)naturalness was difficult. Furthermore, when asked whether some of these contours implied different meanings, these speakers answered that the 'high flat' contour carried an implication of continuation, but that the other contours had no connotations of this kind.

This willingness to accept variation in tonal realisations is difficult to reconcile with an accent-based account, but expected for tones associated with prosodic phrases. The realisation of accents is fixed in that starred tones are by definition aligned with stressed syllables. When a tone is associated with a prosodic domain like the P-phrase, however, it is expected that its location is much more variable. Employing this line of argumentation, Féry (2010) suggests that several Indian languages displaying a small set of tonal targets with variable alignment are best modelled by associations of these tones to the phrase level. She classifies them as part of typological group she calls phrase languages, contrasting with tone languages like Thai and intonation languages like English (see Féry 2015 for more detail). Crucially, phrase languages lack the wealth of intonationally marked pragmatic distinctions that characterise intonation languages (for example, a speaker of English might say *I have some great ideas* in a way that implies that the addressee does

not). The present data support the idea that Finnish is a phrase language as well. As mentioned in the introduction, phonological descriptions of Finnish generally agree that it overwhelmingly displays tonal contours of the same type. It thus lacks the array of contrasting accents that is characteristic of intonation languages (evidently, it lacks lexical tones as well). Instead, Arnhold (2014) presents production data from conditions which usually induce different accents in intonation languages, suggesting that the resulting variation is best accounted for in terms of prosodic phrasing. The connection between the absence of contrasting accents and the flexibility of tonal targets seems natural: In a language like English, displacing the realisation of a starred tone risks leading to the perception of a different accent. When the intonational inventory contains no contrasting accents, only a small set of phrase tones, the absence of competition allows for more variability.

6. CONCLUSION

This article reports a two-alternative forced-choice perception study empirically comparing the differences between three accounts of Finnish intonation employing H_PL_P phrase tones, L+H*, and LHL accents, respectively. While naturally occurring Finnish intonation is characterised by a series of rise-falls, the experimental stimuli included manipulations containing only parts of this contour, in particular the patterns 'rise', 'fall' and 'high fall', in addition to the contours 'flat' and 'rise-fall' approximating natural compound and noun phrase realisations, respectively. Only the 'high fall' manipulation patterned clearly with the 'rise-fall' contour. These results indicate that the realisation of an H and a following L target was necessary for a word to be perceived as carrying a pitch contour of its own, favouring the H_PL_P account. Additionally, the findings support an analysis of the tonal targets as associated with the P-phrase level. Finally, the study provides evidence that Finnish differs from typical intonation languages like English and instead belongs to a different typological group, called phrase languages by Féry (2010, 2015) (see Arnhold 2015a, b for further evidence from production data).

ACKNOWLEDGEMENTS

was supported by Ph.D. grants from the German National Academic Foundation (Studienstiftung des deutschen Volkes) and the Anna Ruths Foundation.

NOTES

1. In broad focus or all-new realisations, finite verbs frequently do not carry a rise-fall contour of their own and it is commonly assumed that they are unaccented in neutral renditions (Välimaa-Blum 1988:99, 1993:84f.; Suomi et al. 2008:114). However, realisations like the one in Figure 1, with finite verbs carrying contours of their own, do appear in these contexts as well (for a quantitative investigation see Arnhold et al. 2010). The matter is not crucial for the topic of this article; the example was chosen purely for ease of exposition.

2. There are few descriptions of Finnish intonation that do not employ the autosegmental-metrical model. Iivonen (1998) provides an account that is explicitly not affiliated with any theory, while Mixdorff et al. (2002) applied the Mixdorff–Fujisaki model of German to a small corpus of Finnish utterances. Vainio et al. (2010) modelled f0 effects of lexical quantity using Xu's (1999, 2005) target approximation model, but see O'Dell (2003:77–79), Suomi (2009) and Arnhold (2014) for an account of quantity that does not assume distinct tonal targets.

3. The Karjalainen database consists of the annual volumes of the Finnish newspaper *Karjalainen* for the years 1991–1997. It contains 34.5 million word tokens and was compiled by Jussi Niemi and his colleagues at the University of Eastern Finland (formerly University of Joensuu). The corpus is part of the Finnish Text Collection, available online through the IT Centre for Science Ltd. (CSC, https://research.csc.fi/-/finnish-text-collection).

4. An alternative analysis is possible: The final fall could be due exclusively to an L_I tone, meaning that L_P was not realised either. Ascribing the final fall to an I-phrase level boundary tone would mean even more clearly that the segmentally ambiguous stretch is realised as only one P-phrase. Thus, it would likewise predict that the 'fall' contour should predominantly induce compound perceptions.

REFERENCES

Arnhold, Anja. 2014. *Finnish Prosody: Studies in Intonation and Phrasing*. Ph.D. dissertation, Goethe-University Frankfurt.

Arnhold, Anja. 2015a. Complex focus marking in Finnish: Expanding the data landscape. Ms., University of Alberta.

Arnhold, Anja. 2015b. Finnish as a phrase language. Ms., University of Alberta.

Arnhold, Anja, Martti Vainio, Antti Suni & Juhani Järvikivi. 2010. Intonation of Finnish verbs. *Speech Prosody 2010*, 100054:1–4.

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4), 390–412.

Bates, Douglas M., Martin Maechler & Ben Bolker. 2011. lme4: Linear mixed-effects models using S4 classes. http://lme4.r-forge.r-project.org/ (accessed 24 July 2011).

Boersma, Paul & David Weenink. 2010. Praat: Doing phonetics by computer. http://www.fon.hum.uva.nl/praat/ (accessed 15 March 2010).

Bruce, Gösta. 1977. *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.

Chen, Aoju. 2011. What's in a rise: Evidence for an off-ramp analysis of Dutch intonation. *17th International Congress of Phonetic Sciences* (ICPhS XVII), 448–451.

Féry, Caroline. 2010. Indian languages as intonational 'phrase languages'. In Imtiaz Hasnain & Shreesh Chaudhury (eds.), *Problematizing Language Studies: Festschrift for Rama Agnihotri*, 288–312. Delhi: Aakar Books.

Féry, Caroline. 2015. *Intonation and Prosodic Structure*. Oxford: Oxford University Press.

Gussenhoven, Carlos. 2000. The phonology of intonation. *Glot International* 6(9/10), 271–284.

Gussenhoven, Carlos. 2008. Semantic judgments as evidence for the intonational structure of Dutch. *Speech Prosody 2008*, 297–300.

Hayes, Bruce & Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language & Linguistic Theory* 9, 47–96.

Iivonen, Antti. 1998. Intonation in Finnish. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation Systems: A Survey of Twenty Languages*, 311–327. Cambridge: Cambridge University Press.

Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4), 434–446.

Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Mixdorff, Hansjörg, Martti Vainio, Stefan Werner & Juhani Järvikivi. 2002. The manifestation of linguistic information in prosodic features of Finnish. *Speech Prosody 2002*, 511–514.

Niemi, Jussi. 1984. *Word Level Stress and Prominence in Finnish and English: Acoustic Experiments on Production and Perception*. Joensuu: University of Joensuu.

O'Dell, Michael. 2003. *Intrinsic Timing and Quantity in Finnish*. Tampere: Tampere University Press.

Pierrehumbert, Janet B. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation, MIT.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/ (accessed 8 July 2011).

Suomi, Kari. 2009. Durational elasticity for accentual purposes in Northern Finnish. *Journal of Phonetics* 37(4), 397–416.

Suomi, Kari, Juhani Toivanen & Riikka Ylitalo. 2008. *Finnish Sound Structure: Phonetics, Phonology, Phonotactics and Prosody*. Oulu: University of Oulu.

Vainio, Martti & Juhani Järvikivi. 2007. Focus in production: Tonal shape, intensity and word order. *Journal of the Acoustical Society of America* 121(2), EL55–EL61.

Vainio, Martti, Juhani Järvikivi, Daniel Aalto & Antti Suni. 2010. Phonetic tone signals phonological quantity and word structure. *Journal of the Acoustic Society of America* 128(3), 1313–1321.

Välimaa-Blum, Riitta. 1988. *Finnish Existential Clauses – Their Syntax, Pragmatics and Intonation*. Ph.D. dissertation, The Ohio State University.

Välimaa-Blum, Riitta. 1993. A pitch accent analysis of intonation in Finnish. *Ural-Altaische Jahrbücher* 12, 82–94.

Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27, 55–105.

Xu, Yi. 2005. Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46(3–4), 220–251.