Proceedings of the ICPhS XVII

17-21 August 2011 Hong Kong

U City University
of Hong Kong

t

Proceedings of
the 17th International
Congress of Phonetic Sciences

17-21 August 2011
Hong Kong

www.icphs2011.hk

Edited by

Wai-Sum Lee & Eric Zee

WHAT'S IN A RISE: EVIDENCE FOR AN OFF-RAMP ANALYSIS OF DUTCH INTONATION

Aoju Chen

Utrecht University; Max Planck Institute for Psycholinguistics, the Netherlands
aoju.chen@mpi.nl

ABSTRACT

Pitch accents are analysed differently in an on-ramp analysis (i.e. ToBI) and an off-ramp analysis (e.g. Transcription of Dutch Intonation - ToDI), two competing approaches in the Autosegmental Metrical tradition. A case in point is pre-final high rise. A pre-final rise is analysed as H* in ToBI but is phonologically ambiguous between H* or H*L (a (rise-)fall) in ToDI. This is because in ToDI, the L tone of a pre-final H*L can be realised in the following unaccented words and both H* and H*L can show up as a high rise in the accented word. To find out whether there is a two-way phonological contrast in pre-final high rises in Dutch, we examined the distribution of phonologically ambiguous high rises (H*(L)) and their phonetic realisation in different information structural conditions (topic vs. focus), compared to phonologically unambiguous H* and H*L. Results showed that there is indeed a H*L vs. H* contrast in prefinal high rises in Dutch and that H*L is realised as H*(L) when sonorant material is limited in the accented word. These findings provide new evidence for an off-ramp analysis of Dutch intonation and have far-reaching implications for analysis of intonation across languages.

Keywords: pre-final rise, off-ramp analysis, on-ramp analysis, intonation, Dutch

1. INTRODUCTION

In intonation languages, pitch accents are analysed differently in ToBI (proposed initially for American English) [3] and ToDI (proposed initially for Dutch) [8], two competing phonological analyses in the Autosegmental-Metrical tradition [11, 13]. ToBI has an on-ramp analysis of pitch accents, according to which the tonal target associated with the stressed syllable and the contour leading to the target are considered the phonological elements of a pitch accent; ToDI has an off-ramp analysis of pitch accents, according to which the tonal target associated with the stressed syllable and the contour in the following unstressed segments (until the next pitch accent) are considered the phonological elements of a pitch accent [9].

A direct consequence of this striking difference between these two analyses is that a pre-final high rise (i.e. either a prenuclear rise or a nuclear rise followed by one or more unaccented words) is simply a high accent (H*) in an on-ramp analysis but is phonologically ambiguous between H* or a (rise-)fall accent (H*L) in an off-ramp analysis. More specifically, in an off-ramp analysis, the L tone or the falling portion of a pre-final H*L is often realised in the following unaccented words. Consequently, both H* and H*L can show up as a high rise in the accented word. This ambiguity can only be resolved in the following (content) word. The high rise is labelled as H* if the following word is either accented or unaccented but with the pitch remaining high; it is labelled as H*L if the following word is unaccented and the pitch falls. In this study, we addressed the question as to whether there is a two-way phonological contrast (H* vs. H*L) in pre-final high rises in Dutch, as claimed in an off-ramp analysis. In doing so, we intended to provide evidence for or against an off-ramp analysis of Dutch intonation.

2. METHOD

In an earlier study on the intonational realisation of topic and focus in Dutch [1], SVO sentences (e.g. *Een poetsvrouw pakt de vaas* 'A cleaning-lady is picking up the vase) were elicited from native speakers of Dutch as answers to either WHO-questions (e.g. *Wie pakt de vaas?* 'who is picking up the vase?) or WHAT-questions in a picture-matching game. In half of the answer sentences, the subject NP was the focus (i.e. the referent required via the WH-word), the object NP was the topic (i.e. the referent introduced early on in the discourse and mentioned again in the WH-question); in the other half of the answer sentences, the subject NP was the topic, the object NP was the focus. Each noun occurred in both the focus condition and the topic condition. Approximately

half of the subject nouns were monosyllabic words and half were bi- or multisyllabic (hereafter multisyllabic) words. The subject nouns were accented in all but four cases. The following accent patterns were observed: (1) a (rise-)fall, (2) a high rise followed by an accented verb, (3) a high rise followed by an unaccented verb with a sustained high pitch, and (4) a high rise followed by a gradual fall covering the unaccented verb, the article preceding the object NP, and in some cases also the object noun. In [1], patterns (2) and (3) were transcribed as H*. Patterns (1) and (4) were transcribed as H*L. The distributions of H* and H*L were similar in the topic and focus conditions.

The data in [1] provided the ideal material to find out whether there is indeed a two-way phonological contrast in pre-final rises in Dutch. For the sake of clarity, we refer to the phonologically ambiguous high rise, i.e. the high rise followed by a gradual fall in the subsequent unaccented words (pattern 4), as H*(L) and the unambiguous (rise-)fall (pattern 1) as H*L. If the H*(L) pattern was a phonetic variant of H*L, we would expect to see the following two outcomes. First, speakers would realise H*L as H*(L) for structural motivation instead of functional motivation because these two patterns should signal identical meanings. The most plausible structural motivation would be that speakers opt to realise the fall after the accented word due to limited sonorant material in the accented word. This led to the hypothesis that H*(L) would occur more frequently in monosyllabic words than in multisyllabic words whereas such a tendency would be weak or absent in the case of H*L and H* (hereafter Hypothesis 1). Further, it was found in [2] that although H*L occurred similarly frequently in sentence-initial topic and focus, the phonetic realisation of H*L (i.e. pitch scaling and peak/valley alignment) differed systematically in the topic and focus conditions. H*L was realised with a wider fall, caused by a lower pitch minimum after the pitch peak, an earlier alignment of pitch maximum, and a later alignment of pitch minimum in the focus condition than in the topic condition. If H*(L) was a phonetic variant of H*L but H* was a different phonological category, we would expect H*(L) to differ phonetically in topic and focus in the same way as H*L but in a different way than H* (hereafter Hypothesis 2).

To test Hypotheses 1 and 2, we recoded the intonational transcription in sentence-initial nouns (N= 280) produced by nine of the Dutch speakers (3 women, 6 man) from [1], and acoustically annotated the high rises in these nouns.

2.1. Recoding high rises

The high-rise versions of H*L (pattern 4) were recoded as H*(L). There were 67 tokens of H*(L) in addition to 154 tokens of H*L (pattern 1), 52 tokens of H* (patterns 2, 3), 4 unaccented nouns, and 3 nouns accented with a infrequent accent type in the current data set.

2.2. Acoustic annotation

The phonetic realisation of H*(L) spanned across the accented noun and the following unaccented portion of the sentence. Considering that it was rather difficult to reliably identify the exact point at which the highest pitch of H*(L) was reached, we focused on acoustic annotation necessary for analysis of pitch scaling. Within the accented noun, pitch scaling of H*(L) could differ along the following parameters: initial pitch-minimum (the lowest pitch preceding pitch-maximum), pitch-maximum, pitch span of the rise (span-rise). Outside the accented noun, pitch scaling of H*(L) could differ in the pitch-minimum in the following unaccented segments (pitch-minimum 2) and the span of the fall from the accented noun to the following unaccented segments (span-fall), comparable to the pitch-minimum after the peak and span-fall within the accented noun in the case of H*L. The phonetic realisation of H* was confined to variation in the accented noun. But for the sake of comparability, H*(L) and H*-accented nouns were annotated for the same acoustic landmarks.

Three pitch-related landmarks were labelled in each of the selected subject nouns in Praat:

- H1: the highest pitch in the noun
- L1: the lowest pitch preceding H1 in the noun
- L2: the lowest pitch in the unaccented verb in the case of H*(L); the lowest pitch in the accented verb in the case of H*

The pitch values (in semitones with 1 Hz as the reference point) at the pitch-related landmarks were automatically extracted by means of Praat scripts. Five measures were obtained from each subject noun:

- Pitch maximum: pitch at the H1 landmark ($pitch_{H1}$)
- Pitch minimum-1 (pitch minimum before pitch maximum): pitch at the L1 landmark ($pitch_{L1}$)
- Pitch minimum-2 (pitch minimum after pitch maximum): pitch at the L2 landmark ($pitch_{L2}$)

- Range-rise: $pitch_{H1}$ - $pitch_{L1}$
- Range-fall: $pitch_{H1}$ - $pitch_{L2}$

It should be mentioned that the verbs were identical in the topic and focus conditions. This made it possible to conduct interpretable comparisons in pitch minimum-2 and span-fall between the two conditions. Differences found could be reliably attributed to the difference in information structure.

3.   ANALYSIS AND RESULTS

3.1.   Distribution of H*, H*(L), and H*L

To assess what determined choice of accent pattern in the subject nouns, we carried out multinomial logistic regression modelling at the significance level of 0.05. The outcome variable was the accent patterns in the nouns, consisting of three categories (H*, H*(L), H*L). The nouns which were either unaccented or accented with a rare accent were grouped into the category OTHER. Due to the very small number of tokens, the OTHER category was not included in the modelling. The predictors were INFORMATION STRUCTURE and SYLLABLE NUMBER; the variable SPEAKER was used to define the subgroups of the data in the model. The H*(L) accent was used as the reference category for pair-wise comparisons.

The model fitting was significantly improved after adding INFORMATION STRUCTURE and SYLLABLE NUMBER as the predictors (% = 81.79, df = 4, p < 0.0001). The likelihood ratio test however showed that only the predictor SYLLABLE NUMBER (x^2 = 81.22, df = 4, p < 0.0001) led to significant improvement in the model fitting. There was thus only a robust relationship between accent patterns and syllable structure. The Wald statistics showed that H*(L) was 5.59 times and 18.87 times more likely to occur in monosyllabic words than in multisyllabic words, compared to H* (Wald = 15.88, df = 1, p < 0.0001, Exp(B) = 0.18) and H*L (Wald = 57.15, df = 1, p < 0.0001, Exp(B) = 0.05) respectively. These results indicated that monosyllabic words attracted H*(L), but to a less extent H* and in particular H*L. In fact, H*L occurred predominantly in multisyllabic words, as can be seen in Table 1.

**Table 1:** Distributions of accent patterns in monosyllabic and multisyllabic subject nouns.

	H*	H*(L)	H*L	OTHER	Total
monosyllabic	25.2%	43.5%	28.2%	3.1%	100%
multisyllabic	13.8%	5.9%	78.3%	2.0%	100%

3.2.   Phonetic realisation of H* and H*(L)

Mixed-effect modelling was used to establish the effect of topic and focus on the five pitch-related measures. For each measure (the dependent variable), a mixed-effect model was built in R with INFORMATION STRUCTURE (i.e. topic, focus) as the fixed-effect factor and SPEAKER and TARGET WORD as the random-effect factors. We report p-values based on the t-statistics as well as the p-values obtained from Markov Chain Monte Carlo (MCMC) sampling with 10000 samples [4].

Regarding H*, a main effect of INFORMATION STRUCTURE (p|MCMC| = 0.06, p|T| = 0.05) was found in the model for span-rise. Span-rise was significantly wider in focus (3.66 st) than in topic (2.66 st). Regarding H*(L), a main effect of INFORMATION STRUCTURE (p|MCMC| = 0.04, p|T| = 0.02) was found in the model for span-fall. Span-fall was significantly wider in focus (9.03 st) than in topic (5.30 st). Further, a main effect of INFORMATION STRUCTURE (p|MCMC| = 0.13, p|T| = 0.004) was found in the model for pitch minimum-2, although the effect was only significant in the t-test. Pitch minimum-2 was significantly lower in focus (79.63 st) than in topic (82.13 st). Stylised realisations of H* and H*(L) in the topic and focus conditions are shown in Figure 1.

Figure 1: Stylised realisations of H* and H*(L) in the topic and focus conditions. The filled square stands for the accented subject noun.
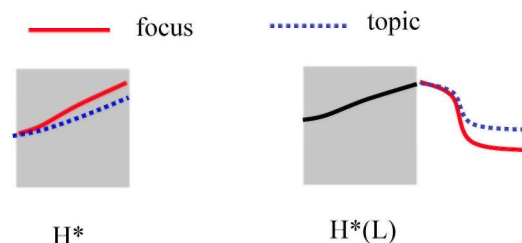
4.   DISCUSSION AND CONCLUSIONS

The analysis on choice of accent pattern in the subject nouns indicates that H*(L) had a stronger preference for monosyllabic words to multisyllabic words, compared to H* and in particular H*L, which occurred predominantly in multisyllabic words. Hypothesis 1 is therefore borne out.

The results from the modelling on H*(L) show that H*(L) was realised with a wider span-fall in the focus condition than in the topic condition, as found for H*L. The difference in span-fall was mainly triggered by a lower pitch minimum after the peak in the focus condition, again as found for

H*L. Further, H*(L) did not differ consistently in span-rise in the two conditions. In contrast, H* was realised with a wider span-rise in the focus condition than in the topic condition but did not differ consistently in span-fall. Our analyses on pitch scaling have thus shown that H*(L) was similar to H*L but different from H* in how phonetic realisation differed in topic and focus. Hypothesis 2 is thus borne out.

Based on these results, we conclude that there is a H*L vs. H* contrast in Dutch prefinal high rises and that H*L is realised as H*(L) when limited sonorant material is available in the accented word. Our study thus provides first evidence for an off-ramp analysis of pre-final rises in Dutch. This adds to the existing evidence for an off-ramp analysis of prenuclear [9] and nuclear rise-falls [10], supporting the appropriateness of an off-ramp analysis for Dutch intonation.

Our results have far-reaching implications for analysis of intonation across languages in the Autosegmental-Metrical framework. In recent years, a wide range of languages have adopted an on-ramp analysis of intonation (e.g. Catalan, German, Italian, Spanish etc.). Other than ToDI, an off-ramp analysis has been proposed for British English, European Portuguese, and Russian [5, 6, 12]. It is however not always clear why a language adopts one but not the other analysis of intonation. The limited evidence available in languages other than Dutch is primarily obtained from perceptual studies [7, 14]; the corroborating production evidence is still lacking. The difference between an off-ramp analysis and an on-ramp analysis has been claimed to be of a theoretical nature rather than a typological one [7]. In the light of the limited evidence for an on-ramp analysis of German intonation [7] and Catalan intonation [14] and increasing evidence for an off-ramp analysis of Dutch intonation, we argue that the difference between an off-ramp analysis and an on-ramp analysis is not just a theoretical debate but is typologically real. More specifically, languages can differ in what constitutes a pitch accent or which part of the pitch contour conveys meaning differences. As a result, the intonation of a language can be more appropriately described by adopting an on-ramp analysis, an off-ramp analysis or even both. Examining the phonetic realisation of a pitch accent in different meaning contexts can be a fruitful method to find out what is the most suitable intonational analysis in a certain language, as shown by the present study.

6. REFERENCES

[I]   Baayen, R.H., Davidson, D.J., Bates, D.M. 2008. Mixed-effects modelling with crossed random effects for subjects and items. *J. of Mem. and Lang* 59, 390-412.

[2]   Beckman, M.E., Hirschberg, J., Shattuck-Hufnagel, S. 2005. The original ToBI system and the evolution of the ToBI framework. In Jun, S. (ed.), *Prosodic Typology*. Oxford: OUP, 9-54.

[3]   Chen, A. 2009. The phonetics of sentence-initial topic and focus in adult and child Dutch. In Vigário, M., Frota, S., Freitas, M.J. (eds.), *Phonetics and Phonology: Interactions and Interrelations*. Amsterdam: Benjamins, 91-106.

[4]   Chen, A., Fikkert, P. 2007. Intonation of early two-word utterances in Dutch. *Proc. of the 16th ICPhS* Dudweiler: Pirott, 315-320.

[5]   Frota, S. In press. The intonational phonology of European Portuguese. In Jun, S. (ed.), *Prosodic Typology* 2. Oxford: OUP.

[6]   Grabe, E. 2004. Intonational variation in urban dialects of English spoken in the British Isles. In Gilles, P., Peters, J. (eds.), *Regional Variation in Intonation*. Tuebingen: Niemeyer, 9-31.

[7]   Grice, M., Baumann, S., Benzmüller, R. 2005. German intonation in autosegmental-metrical phonology. In: Jun, S. (ed), *Prosodic Typology*. Oxford: OUP, 55-83.

[8]   Gussenhoven, C. 2005. Transcription of Dutch intonation. In Jun, S. (ed.), *Prosodic Typology*. Oxford: OUP, 118-145.

[9]   Gussenhoven, C. 2008. Semantic judgments as evidence for the intonational structure of Dutch. *Proc. of Speech Prosody* Campinas, 297-300.

[10] Hanssen, J., Jörg P., Gussenhoven, C. 2008. Prosodic effect of focus in Dutch declaratives. *Proc. of Speech Prosody* Campinas, 609-612.

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

[12] Odé, C. 2008. Transcription of Russian intonation. *Dutch Contributions to the 14th International Congress of Slavists,* New York: Rodopi, 431-449.

[13] Pierrehumbert, J.B. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation, MIT.

[14] Prieto, P. In press. The intonational phonology of Catalan. In Jun, S. (ed.), *Prosodic Typology* 2. Oxford: OUP.