



Dutch Sentence Intonation Revisited

Philippe Martin

LLF, UFRL, Université de Paris, France

philippe.martin@linguist.univ-paris-diderot.fr

Abstract

Since the first descriptions of Dutch sentence intonation highlighting their characteristic melodic hat patterns [1], most studies in this field were undertaken using the Autosegmental-Metrical (AM) framework and the ToDI notation system, an optimized ToBI variant for Dutch [2].

In the AM approach, accent phrases are minimal prosodic units (AP), assembled to form Intonation Phrases (IP), and sequences of IPs' constitute the Prosodic Structure (PS). AP's are characterized by pitch accents which do not interact with each other, while IP's are ended by boundary tones.

From the analysis of a selection of read speech recordings sampled from the Corpus Gesproken Nederlands (CGN) [3], an alternate analysis is introduced, in which pitch accents do interact with each other and indicate through dependency relations partial accent phrase hierarchical structures internal to IP's.

Furthermore, instead of ToDI notation, prosodic events are described as melodic contours, above or below a glissando threshold [4], integrating pitch perception in the model. IP's final pitch accent is merged with its boundary tone as a single prosodic event.

Index Terms: Dutch, prosodic structure, ToDI, glissando threshold, melodic contours.

1. Introduction

One of the first experimental analysis of Dutch sentence intonation was reported by 't Hart, Collier and Cohen [1]. Through validation of vocoder stylized pitch pattern perception and fundamental frequency acoustic analysis, they brought to light a characteristic hat pattern fundamental frequency.

Later, in the wave of the Autosegmental-Metrical (AM) model applied to prosodic events [2], prosodic descriptions based on the acoustic analysis of (short) sentences [5, 6] lead to the development of the intonation transcription system ToDI, variant of the well-known ToBI system originally proposed for American English [7]. Numerous detailed studies using this notation system on Dutch data appeared later [8-14], describing specific phonetic and phonological characteristics of Dutch sentence intonation, such as falling-rising pitch contours, focus in declaratives, etc.

The aim of these studies is to establish a grammar of Dutch intonation, resulting from a collection of well-formed intonative prosodic events transcribed with the ToDI notation. Such a grammar assumes that minimal prosodic units are accent phrases (AP), containing only one non-emphatic stressed syllable. As, contrary to French or Korean, Dutch is a lexically stressed language, each accent phrase usually contains one verb, one adverb, one adjective or one noun, i.e. categories of words with one stressed syllable.

In the AM view, the sentence prosodic structure results from a hierarchical assembly of AP's to form Intonational Phrases (IP's) whose sequence constitute the sentence prosodic structure (PS). Prosodic events located on stressed syllables, one per AP, are called pitch accents. Those located on IP's left and right boundaries are boundary tones. In this model, pitch accents do not interact with each other, and the PS is actually defined by boundary tones (the left boundary being optional [14]). According to Gussenhoven [2], there is no intermediate intonation phrase (ip) in Dutch that would assemble AP's in an intermediate level of the prosodic structure.

2. Another Model

Instead of analyzing a new set of data using the AM theoretical approach and the ToDI notation system, a somewhat different analysis model has been applied to a number of read speech recordings in order to evaluate its effectiveness compared to the existing studies on Dutch sentence intonation.

Starting with the same definition of accent phrases as prosodic minimal units, the prosodic structure is then defined as a hierarchical assembly of AP's into multiple levels, the highest one being the Intonation Phrases IP. In this AP's hierarchy forming the sentence prosodic structure, only IP's are possibly indicated by boundary tones. The other levels are marked by specific pitch accents, possibly acting together with boundary tones at the IP level. In this view, pitch accents do interact with each other, depending on the prosodic structure complexity in terms of number of levels and the number of AP's merged together at each level.

Another important difference pertains to the transcription of prosodic events, and particularly of pitch accents and boundary tones. Instead of considering tone targets as in the ToDI system, pitch events are transcribed as melodic contours characterized by a perception criterium based on their glissando value. Glissando values are given by the formula $(st2-st1)/(t2-t1)$ with $st1$ and $st2$ are respectively the semitones values at the beginning and at the end of the pitch change, approximated as linear, and $t2-t1$ is the duration of the pitch movement. Semitones are given by $st = 12 * \log(F0/100.0) / \log(2)$, $F0$ being the fundamental frequency. Values above a glissando threshold equal to $k / (t2-t1)^2$ with $0.16 < k < 0.32$ are assumed to be perceived as melodic changes, those below as static tones at $2/3$ of the pitch change. Compensation for concordant intensity change is also provided [4, 15]. Being a rough estimation, the integration of the glissando value in the annotation of melodic contours constitutes nevertheless a first approximation aiming to integrate perception in the description of melodic contours.

From these considerations, classes of melodic contours are defined a priori from the melodic movement located on the AP's stressed vowel (and final vowel at the IP level) as follows:

- C_{dec} terminal conclusive *déclarative*, falling and reaching the lowest pitch value in the sentence, usually above the glissando threshold. The terminal conclusive interrogative counterpart C_{int} , rising and reaching the highest pitch level in the sentence, above the glissando threshold.
- C_{ris} rising contour above the glissando threshold.
- C_{fal} falling contour above the glissando threshold.
- C_{neu} neutralized contour rising or falling but below the glissando threshold.
- C_c complex contour, falling on the AP stressed vowel and rising on the final vowel and usually followed by a pause. All other contours are located on the AP's stressed vowel.

3. Data

The analyzed corpus consists of more than 20 recordings sampled at random from a set of 565 read sentences part of the Corpus of Gesproken Nederlands (CGN) for the Dutch variety [3]. Read sentences were selected according to their apparent increasing complexity based on the number of words, and therefore of accent phrases, they contain.

Acoustic analysis was carried out with a speech analysis software dedicated to fundamental frequency analysis [16], judged preferable to the de facto standard Praat [17] for its prosodic graphic annotation functions and much better F0 tracking capabilities among other features.

The reliability of F0 curve were checked visually by displaying an underlying narrow band spectrogram whose frequency scale is aligned on the fundamental frequency scale, as shown on Fig. 1.

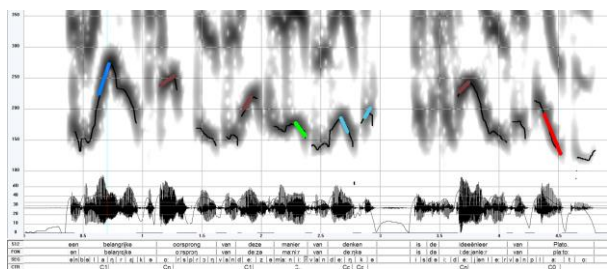


Figure 1: *Aligning the spectrogram and F0 frequency scales on superposed display allows an easy visual validation in case of doubtful reliability of the fundamental frequency curve.*

User friendly graphical commands allows to draw linear segments to approximate sections of F0, corresponding to AP's stressed and IP's final vowels. These segments are automatically labelled and color coded according to the definition of melodic contours given above: in this paper, final terminal C_{dec} are displayed in red, C_{int} in pink, C_c in turquoise, C_{ris} in blue, C_{fal} in green and C_{neu} in brown (colors are user defined), following their glissando values compared to the glissando threshold (evaluated with the 0.16 coefficient). All contours F0 and time values of these contours can be transferred in one mouse click to a spreadsheet such as Excel® together with the corresponding intensity values.

The CGN companion files of each sound file, originally encoded in "awd" format (a variation of Praat TextGrid) using a proprietary SAMPA like phonetic coding, are automatically converted into standard IPA. Occasional errors in CGN original segmentation were corrected securing a proper alignment between stressed and final AP's vowels with the corresponding

F0 curve. Given a reliable segmentation and F0 tracking method, an automatic process for melodic contour annotation can be applied from stressed vowels location, but it has not been used here given the various recording qualities possibly affecting the reliability of F0 curves.

4. Distribution of contours and prosodic structure complexity

The strategy to discover possible regularity in the distribution of melodic contours is based on the analysis of sentences of increasing complexity in terms of accent phrases, and not directly in terms of any syntactic complexity or property. Only a few examples among few hundreds and limited to declarative cases are given here given the limited space available.

4.1. Two accent phrases

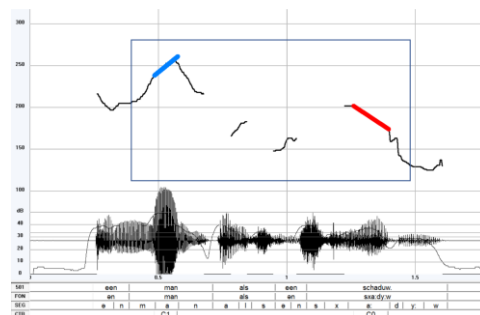


Figure 2: [*een man* C_{ris}] [*als een schaduw* C_{dec}] "a man like a shadow". CGN fn001001 recording.

Fig. 2 shows a first case with only two accent phrases [*een man* C_{ris}] [*als een schaduw* C_{dec}] "a man like a shadow". Accent phrases are specified in squared brackets and stressed vowels are marked in bold underlined characters. The first stressed vowel carries a rising contour C_{ris} above the glissando threshold aligned on the stressed vowel of the *een man* AP. The second AP stressed vowel of *als een schaduw* carries a C_{dec} terminal conclusive contour. The four annotation tiers show respectively the segmentation in words, orthographic and IPA transcription, the segmentation in phones and in melodic contours.

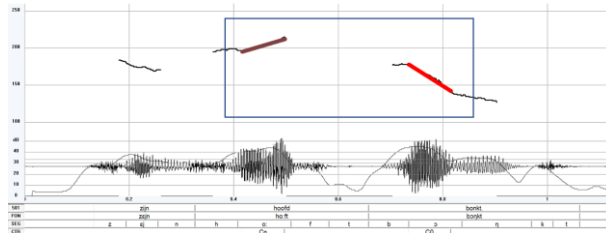


Figure 3: [*zijn hoofd* C_{neu}] [*bankt* C_{dec}] "His head is pounding". CGN fn001001 recording.

Fig. 3 is a variant of the two AP's prosodic structure, illustrating the neutralization of the first melodic contour C_{neu} below the glissando threshold. Indeed, the first accent phrase contour has to be differentiated only from another terminal contour C_{dec} or C_{int} , contours which would encode two PS and two independent sentences. Therefore, this first contour can be realized as a C_{ris} or a C_{neu} , excluding the rarely attested falling contour C_{fal} that could be confused with the terminal contour C_{dec} .

4.2. Four accent phrases

Fig. 4 shows a prosodic structure assembling four AP's as follows: [[*een fietsbel* C_{neu}] [*rinkel* C_{ris}]] [[*in de avond* C_{neu}] [*schemering* C_{dec}]], not congruent with syntax, but more eurhythmic than a congruent version, which would be [*een fietsbel*] [*rinkel* *in de avond* *schemering*]. Although *avondschemering* is orthographically in one word, it possesses two stressed syllables and therefore determines two accent phrases. As in the example of Fig. 3, the first and third contours are neutralized into C_{neu}.

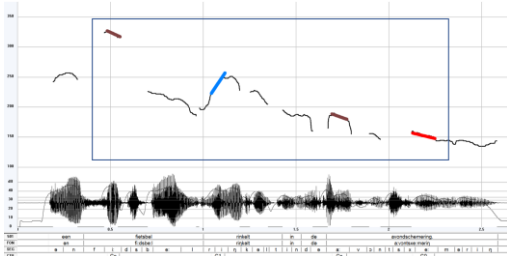


Figure 4: [[*een fietsbel* C_{neu}] [*rinkel* C_{ris}]] [[*in de avond* C_{neu}] [*schemering* C_{dec}]] “a bicycle bell rings in the evening twilight lit”. CGN fn001001 recording.

The next figure (Fig. 5) gives another example of a symmetric expansion in 2 levels of [[*dan pakt* C_{neu}] [*ze altijd* C_{ris}]] [[*en kop* C_{neu}] [*koffie* C_{dec}]], followed with a theme sequence in a Rheme-Theme organization of the sentence characterized by neutralized contours C_{neu} located on each stressed vowel.

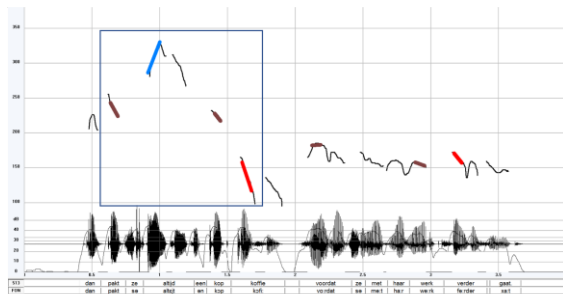


Figure 5: [[*dan pakt* C_{neu}] [*ze altijd* C_{ris}]] [[*en kop* C_{neu}] [*koffie* C_{dec}]] [[*voordat* C_{neu}] [*ze met haar werk* C_{neu}] [*verder gaat* C_{dec}]] “Then she always takes a cup of coffee so that she continues her work”. CGN fn001020 recording.

4.3. A flat sub-prosodic structure

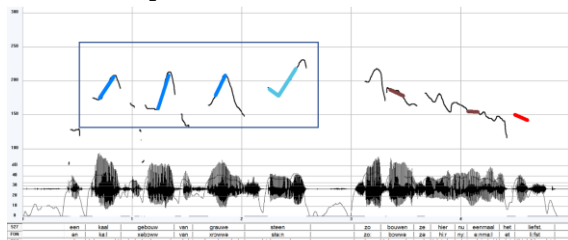


Figure 6: [[*Een kaal* C_{ris}] [*gebouw* C_{ris}] [*van grauwe* C_{ris}] [*steen* C_{con}]] [*zo bouwen* C_{neu}] [*ze hier nu éénmaal* C_{neu}] [*het liefts* C_{dec}]] “They simply love to build a bare gray-stone building here”. CGN fn001019 recording.

Figure 6 illustrates the contrast between Cris rising contour and the complex contour C_c to encode the partial structure

[*Een kaal* C_{ris}] [*gebouw* C_{ris}] [*van Grauwe* C_{ris}] [*steen* C_c] enunciated by the speaker as a list of AP's not congruent with syntax, and followed by neutralized contours C_{neu} before the C_{dec} terminal conclusive. To ensure a proper contrast with the preceding contours C_{ris} a complex contour C_{con} is realized to mark IP boundary.

4.4. Enumeration pattern

The falling contour C_{fal} is mostly found after C_{ris} as a marker of a tight relation between the AP's involved, as found for example, in Fig. 7, in the group [*beste* C_{ris} *resultaat* C_{fal}]. Also, the three first AP's are forming a list with the same C_{ris} contours.

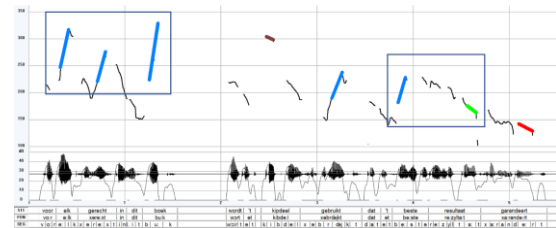


Figure 7: [[*Voor elk* C_{ris}] [*gerecht* C_{ris}] [*in dit boek* C_{ris}] [*wordt 't kipdeel* C_{neu}] [*gebruikt* C_{ris}]] [[*dat 't beste* C_{ris}] [*resultaat* C_{fal}]] [*gerandeert* C_{dec}]] “Each dish in this book uses the chicken part that guarantees the best result”. CGN fn001015 recording.

4.5. Rise-Fall pattern

Another example of the rise-fall C_{ris} C_{fal} configuration is given Figure 8. The two accent phrases involved are *paar* and *minuten*, indicating a close relationship between words, similar to those found in compound words.

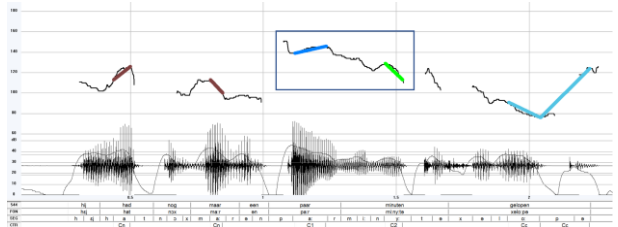


Figure 8: [[*Hij had* C_{neu}] [*nog meer* C_{neu}] [*een paar* C_{ris}] [*minuten* C_{fal}]] [*gelopen* C_{con}]]... “He had walked a few minutes more”. CGN fn001044 recording.

4.6. Rise-Neutral-Fall pattern

The rise-fall configuration C_{ris} C_{fal} can be expanded into 2 levels as C_{ris} C_{neu} C_{fal} (Figure 9), indicating a two-level hierarchy [[*een ascetische* C_{ris}] [*joodse* C_{neu}] [*groep* C_{fal}]].

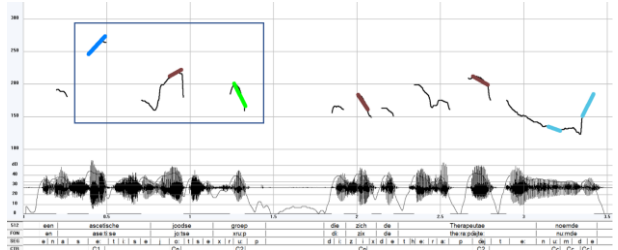


Figure 9: [[*een ascetische* C_{ris}] [*joodse* C_{neu}] [*groep* C_{fal}]] [[*die zich* C_{neu}] [*de Therapeutae* C_{neu}] [*noemde* C_c] ... “an ascetic Jewish group who called themselves the Therapeutae”

4.7. Saturation pattern

Fig. 10 is an example where the melodic contour system runs out of feature contrasts to be congruent with syntax, leaving the realization of stressed syllables to C_{neu} for [oorsprong] and [van deze] AP's: [een belangrijke oorsprong van deze manier van denken]

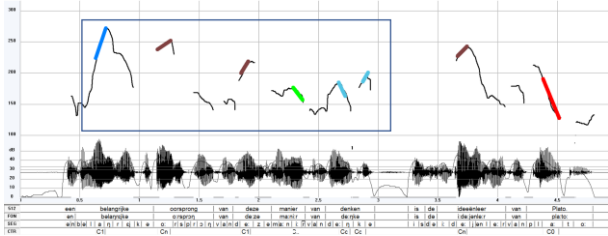


Figure 10: A case of saturation of the prosodic expansion for the group [een belangrijke oorsprong van deze manier van denken] is de ideenleer van Plato “an important origin of this way of thinking is Plato's theory of idea”. CGN fn001019.

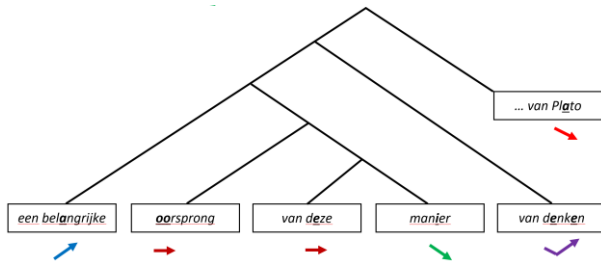


Figure 11: The partial PS for the example of Figure 10, non-congruent to syntax due to the exhaustion of available feature contrasts, leading to the use of a neutralized contour for [oorsprong] and [van deze] AP's. However, a supplementary feature +/- High could be considered to differentiate the two neutralized contours, both below the glissando threshold.

4.8. More complex sentences

The next example Figure 12 shows a more complex sentence organized into 3 IP's, the two firsts ended with C_{con} and the last one by C_{dec} : [er zijn zoveel verschillende manieren om kipfilets te bereiden C_{con}] [dat ze problemloos meerdere keren per week op het menu kunnen staan C_{con}] [zonder dat ze eentonig voorspelbaar worden C_{dec}]. The two first IP's contain a two levels sub-prosodic structure indicated by the sequence of contours C_{ris} C_{neu} C_{fal} : [zoveel C_{ris}] [verschillende C_{neu}] [manieren C_{fal}] and [meerdere C_{ris}] [keren C_{neu}] [per week C_{fal}].

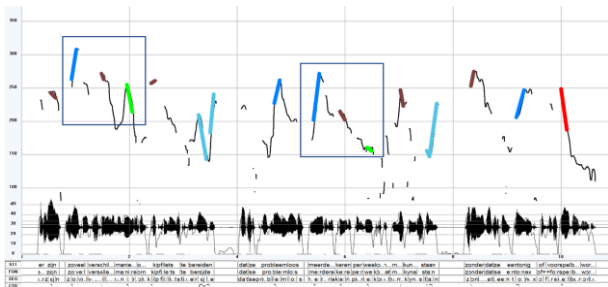


Figure 12. A more complex sentence with the first two IP's are ended by C_{con} and the two last IP's by C_{ris} .

4.9. Minimal and sufficient feature contrasts

Fig. 13 illustrates an application of the minimal and sufficient contrasts principle. The two first IP's [kipfilets he**bben en grote voedingswaarde] [zijn op allerlei manieren toe te passen] contain AP's with C_{ris} so that their final contour must be complex to ensure a proper contrast. However, the two next IP's [zijn voordelig] and [makkelijk] [te bereiden]] can only have a C_{ris} which is sufficient, the first one with a single AP and the second IP's having only a neutralized contour.**

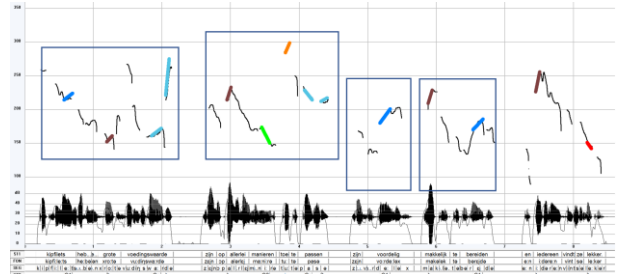


Figure 13: [kipfilets he**bben en grote voedingswaarde] [zijn op allerlei manieren toe te passen] [zijn voordelig] [makkelijk te bereiden]] [en iedereen vindt ze lekker] “chicken fillets have a high nutritional value / can be used in all sorts of ways / are inexpensive to prepare and everyone likes them”. CGN fn001015.**

5. Discussion

Some points of the model presented here may be sometimes subject to interpretation in their application:

1. Melodic contours are aligned on vowels which may be problematic to validate as belonging to syllables known to be effectively stressed. This is frequently the case with compound words, such as *kipfilets* or *kipfilets?* *avondschemering* or *avondschemering?*

2. The glissando threshold determines the class of rising or falling contours as neutralized, i.e. not perceived as a melodic change. However, in the literature, this threshold was established from perception tests conducted on a unique synthesized vowel [a]. The coefficient 0.16 used in the formula vary sometimes up to 0.32 to take this uncertainty into account [15].

Other than that, the role of accent phrase pitch accent as marker of an Intonation Phrase sub-structure appears clearly, once they are described not as tone targets but as melodic contours.

6. Conclusion

An alternate phonological analysis of Dutch intonation is presented, using an approach totally different from the current autosegmental-metrical model using ToDI notation:

1. Prosodic events located on stressed vowels are described by melodic contours characterized by their glissando value and not by tonal targets.
2. Melodic contours define dependency relations between accent phrases, whose grammar define the prosodic structure eventually in more than one level.

Among other benefits, this alternate model provides coherent explanations for melodic contour neutralization in the framework of a prosodic grammar involving dependency relations totally independent from syntax.

7. References

- [1] J. 't Hart, R. Collier & A. Cohen, A perceptual study of intonation: an experimental-phonetic approach to speech melody. Cambridge: Cambridge University Press, 1990. pp. xv + 212.
- [2] C. Gussenhoven, Transcription of Dutch Intonation, in "Prosodic Typology: The Phonology of Intonation and Phrasing", Sun-Ah Jun ed., Oxford University Press, 2005.
- [3] Corpus of Gesproken Nederlands: <http://lands.let.ru.nl/cgn/>
- [4] M. Rossi, "Le seuil de glissando ou seuil de perception des variations tonales pour la parole", *Phonetica* (23), 1971, pp.1-33.
- [5] C. Gussenhoven, and T. Riad *Tones and Tunes*. Volume 1. pp. 385, Berlin, New York, 2007.
- [6] T. Riad and C. Gussenhoven *Tones and Tunes*. Volume 2 pp. 415, Berlin, New York, 2007.
- [7] Beckman, M. E., Hirschberg, J. & Shattuck-Hufnagel, S. "The original ToBI system and the evolution of the ToBI framework", In S.-A. Jun (ed.), *Prosodic typology – The phonology of intonation and phrasing*. Oxford University Press, 2005, pp. 9-54.
- [8] S A Jun and C. Gussenhoven *Prosodic Phonology II* pp. 324 - 364, Oxford UK, 2014.
- [9] C. Gussenhoven, J. Peters and J. Hanssen "Alignment of the second low target in Dutch falling-rising pitch contours", *Proceedings of Interspeech 2007*, pp. 2641- 644.
- [10] J Hanssen, C. Gussenhoven and J Peters "Phrase-final pitch accommodation effects in Dutch", *Proceedings of the 16th International Congress of Phonetic Sciences*, 2017, pp. 1077-1080.
- [11] C Gussenhoven, J Peters and J Hanssen "Prosodic effects of focus in Dutch declaratives" *Proceedings of Speech Prosody* pp. 609 – 612, 2008.
- [12] M. van de Ven and C. Gussenhoven "The timing of the final rise in falling-rising intonation contours in Dutch", *Journal of Phonetics* pp. 225-236, 2011.
- [13] J Peters, J Hanssen and C Gussenhoven "The timing of nuclear falls", *Laboratory Phonology* pp. 1-52, 2015.
- [14] Intonatie, <http://www.taalportaal.org/taalportaal/topic/pid/topic-14020545824239503>.
- [15] P. Mertens, « Un outil pour la transcription de la prosodie dans les corpus oraux », *Traitement Automatique des langues* 45 (2), 2004, 109-130.
- [16] Ph. Martin, WinPitch: www.winpitch.com
- [17] P. Boersma and D. Weenick, Praat: www.praat.org