

Mandarin-accented Fall, Rise and Fall-Rise F0 contours in Dutch

Xuliang He^{1,2}, Judith Hanssen², Vincent J. van Heuven³ and Carlos Gussenhoven^{2,4}

¹School of Foreign Studies, Nantong University

²Centre for Language and Communication, Radboud University Nijmegen

³Leiden University Centre for Linguistics

⁴Queen Mary University of London

{x.he, j.hanssen,c.gussenhoven}@let.ru.nl}, v.j.j.p.van.heuven@hum.leidenuniv.nl

Abstract

An experiment was conducted to study how Mandarin Chinese speakers of Dutch produced three Dutch pitch contours (Fall, Rise and Fall-rise) on final syllables in the intonational phrase. It was found that the Fall-Rise was the hardest contour for them to produce. In addition, there were substantial differences between the subjects with higher and those with lower proficiency in Dutch in the realization of the Fall and the Rise, which the higher proficiency subjects produced much as did the native control group.

Index Terms: Mandarin speakers of Dutch, Mandarin, Dutch, prosody, intonation contours

1. Introduction

Dutch is an intonation-only language which distinguishes nuclear Fall, Rise, and Fall-rise intonation contours, besides a large number of other nuclear contours [7], while Mandarin is a tone language with four lexical monosyllabic pitch contours. The Dutch intonation contours will appear on single syllables whenever the nuclear accent occurs on the last syllable of the intonational phrase (IP). While in Mandarin a Rise occurs as Tone 2 and a Fall as Tone 4 [11], the Fall-Rise, which is carried out entirely in the syllable rime, does not have a Mandarin counterpart (Mandarin Tone 3 begins any pitch fall earlier and lower than the Fall-rise.) There are hardly any studies on how Chinese learners produce and use the intonation contours of Germanic languages. Hong Kong English speakers used more rises in different types of conversations with native speakers than other contours according to [12]. Chinese learners of English sometimes substituted Falls and Rises for the level tone, especially those with lower English proficiency [2]. Overall, their pitch span was narrower than that of native British English speakers.

Generally, simple tones are acquired before complex tones, and falls are produced before rises [6, 8]. So the monosyllabic Fall-rise of English and Dutch should be expected to be difficult for Chinese learners of these languages to acquire, not only because in Chinese there is no equivalent contour in the tone inventory, but also because it is a complex tone. The availability of phonetic space for the realization of an intonation contour contributes to phonetic variation. Speakers can adjust the phonetic shape of a nuclear pitch contour either by producing a reduced version of the contour or by increasing the speed of the pitch movement in order to produce the full contour. The first is known as ‘truncation’, the latter as ‘rate adjustment’ [4], renamed ‘compression’ by [3]. The adjustment strategies to diminished availability of phonetic space were found to be cross-linguistically different by [5].

The Dutch rising-falling peak, equivalent to the Fall, was found to be time-compressed rather than frequency-compressed by [1]. First, the slopes of the rise and fall parts were steeper in shorter vowels than in longer vowels. Second, time pressure induced by tempo increases greatly compressed the

accent-lending rise but not the fall. The onset of the rise aligned with the onset of the syllable, but the alignment of its end varied with the beginning of the fall. As a result, neither the beginning nor the end of the fall had a fixed anchor point in the segmental string. The effect of the availability of voiced material on the realization of the three nuclear contours in Standard Dutch as used on three types of sentences was investigated by [9]. The choice of contour had different effects on the test words. The Fall and the Rise were somewhat compressed as well as somewhat truncated, and the overall pitch range was reduced as sonorant portions were shorter. By contrast, only the pitch range was compressed in the Fall-rise. Moreover, [10] found that the Fall-rise on IP-final syllables in Dutch was less frequently used than on IP-medial nuclear syllables.

In the experiment reported here, we tried to answer the following questions:

I. Can Mandarin Chinese speakers of Dutch produce the target contours on IP-final monosyllabic words and to what extent are their realizations different from those by native speakers of Dutch (NSD)?

II. Do adjustment strategies to diminished availability of phonetic space adopted by the NSD differ from those used by Mandarin Chinese speakers of Dutch with higher proficiency (CHD) and with lower proficiency (CLD)?

2. Method

2.1. Materials

To create time pressure, we adopted the methodology of [9]. Four test words with increasing durations of the sonorant rime (*Lof* [lɔf], *Loof* [lo:f], *Lom* [lɔm], *Loom* [lo:m]) were put into short carrier sentences intended to elicit three different nuclear contours, the Fall (H*L L%), the (low) Rise (L*H H%) and the Fall-rise (H*L H%) [7], exemplified in (1) to (3), respectively. The rhythmic structure and the number of syllables before the test word was held constant and the new information was located in the final target word. All carrier sentences were preceded by a context sentence, with which they formed mini-dialogues. Two example mini-dialogues with each of the three intonation contours that were to be elicited were provided to the subjects before each of the three blocks of intonation contours. Besides the two practice mini-dialogues, each block contained four experimental mini-dialogues, which always differed from the practice mini-dialogues.

(1) Statements

A: *Met wie gaat je baas morgen trouwen?*

‘Who is your boss going to marry tomorrow?’

B: *Hij trouwt met mevrouw de Loom.*

‘He is going to marry Mrs de Loom.’

(2) Yes/No questions

A: *Ik moet straks naar de baas komen, omdat ik weer te laat was vanmorgen.*

‘I’ll have to see the boss in a minute, because I was late again this morning.’

B: *Moet je naar dokter Loom? Oei, maak je borst dan maar nat!*

‘Do you have to go see doctor Loom? Goodness, then you can expect the worst!’

(3) Rhetorical questions

A: *Dit antieke horloge is nog van opa Thijssen geweest.*

‘This antique wristwatch belonged to grandfather Thijssen’

B: *Het was toch van opa Loom?*

‘Wasn’t it grandfather Loom’s?’

2.2. Procedure

The practice mini-dialogues and the context sentences were recorded by two native speakers in the studio of the Arts Faculty of Radboud University Nijmegen. The text of the practice mini-dialogues and of the randomized experimental mini-dialogues for each type of intonation contour were presented on screen to subjects, together with all sound files as spoken by the two native speakers. Subjects were asked to practise the response sentence in the example dialogues in each block by listening to the sound files before the recording took place. If they could do this, as judged by the first author, the recording began. Subjects were recorded in different locations with a Zoom H4 recorder (48 kHz, 16 bit). They were allowed to repeat any sentence as often as they wished. A group of 20 Chinese speakers of Dutch (3 male), aged from 17 to 53, participated in the production experiment. They had been divided into a higher (CHD) and a lower subgroup (CLD) on the basis of their mean segmental and prosodic proficiency scores as judged by three experts in an earlier experiment. There were 23 subjects in the native control group (9 male), aged from 14 to 49.

2.3. Acoustic measurements

We annotated the data as in Figure 1, following [9].

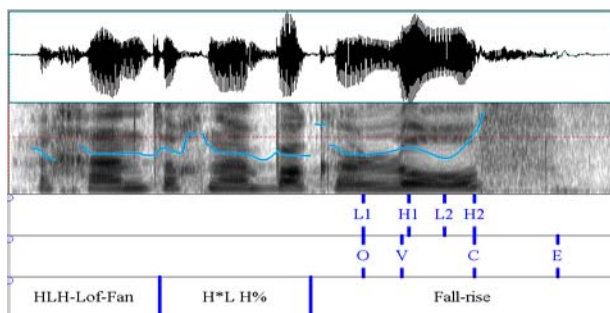


Figure 1. *Tonal and segmental labels on a target word in We gaan toch naar Bakker Lof? read by a CSD as established on the basis of auditory evaluation as well as visual inspection of the waveform, the pitch track and the spectrogram.*

The tonal labels of the f_0 minima and maxima of the three nuclear contours were placed on the first tier (L1: f_0 elbow preceding nuclear peak; H1: Nuclear f_0 peak; L2: End of nuclear fall; Elbow between two f_0 maxima in fall-rise; H2: End of nuclear in fall-rise). The segmental boundaries of the test words were labeled on the second tier and were placed at positive zero-crossings (O: Beginning of onset of accented word; V: Beginning of nucleus of accented word; C: Beginning of coda of accented word; E: End of coda of accented word). The test sentence, the nuclear contour and any comments were provided on the third. We measured the following variables:

- Duration (in ms) of the sonorant rime
- Duration (in ms) of the Fall, Rise and Fall-rise
- Distance (in ms) of f_0 maximum to beginning of the vowel
- Location (in %) of f_0 maximum relative to vowel duration
- Excursion (in semitones, st) of the Fall, Rise and Fall-rise
- The speed (in st/s) of the Fall, Rise and Fall-rise
- F0 (in st) at the labels.

3. Analysis

Of the 80 utterances per contour produced by the Chinese speakers of Dutch, the percentages with correct intonation and accentuation of the target words are 90 (Fall), 100 (Rise) and 73 (Fall-rise). The Fall-rise was evidently the most difficult pitch configuration for these learners, who had few problems with pronouncing the Fall and the Rise.

3.1. SRD (Sonorant rime duration) comparisons

The SRDs between the three groups are significantly different in Fall [$F(2, 35) = 6.6, p < .05, \eta^2 = .27$], Rise [$F(2, 38) = 6.6, p < .05, \eta^2 = .26$] and Fall-rise [$F(2, 38) = 34.2, p < .05, \eta^2 = .64$]. Multiple comparisons show that SRD in the three nuclear contours are significantly longer for CLD than for NSD. In Falls, SRD is also significantly longer for CHD than NSD, but there is no significant difference between CLD and CHD. In Rises, there are no significant differences between any two groups. In Fall-rises, the SRD by CLD and CHDs are significantly longer than those by NSD.

In Figures 2, 3 and 4, the positions of L1, H1, L2, H2 and onset of the vowels are indicated by plot symbols for the relevant contours by NSD, CHD and CLD. The timings of H1, H2 and L2 are expressed (in ms) relative to the timing of L1, which was set to zero. L2 in the Fall and L1 in the Rise and Fall-rise were chosen as the reference points to normalize the pitch values of the other labels. The figures were drawn from averaged data of correct accents and nuclear pitch contours.

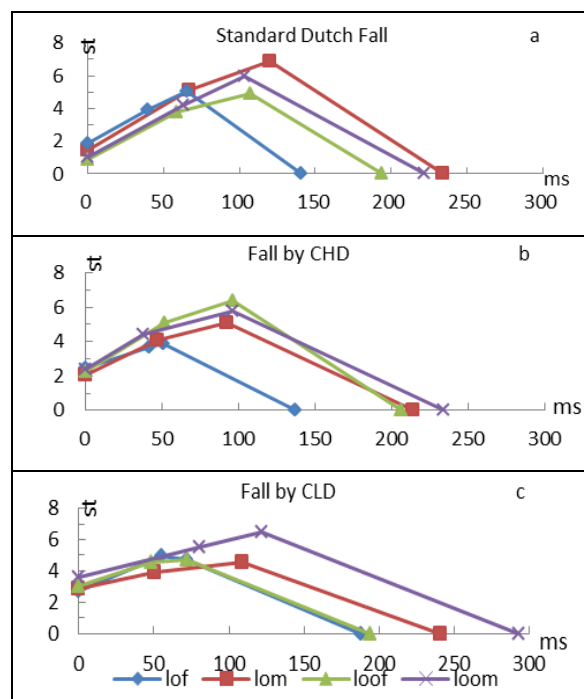


Figure 2: *Falls by native speakers (NSD) and Chinese speakers with higher (CHD) and lower (CLD) proficiency*

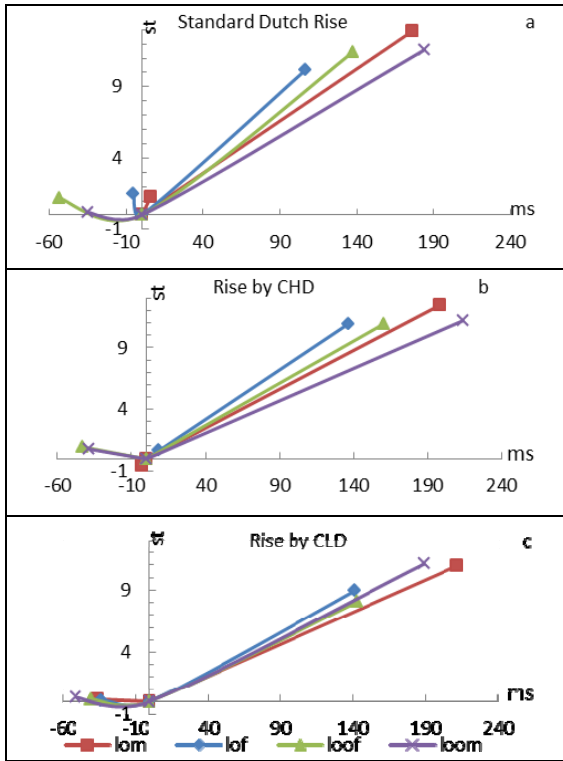


Figure 3: Rises by native speakers (NSD) and Chinese speakers with higher (CHD) and lower (CLD) proficiency.

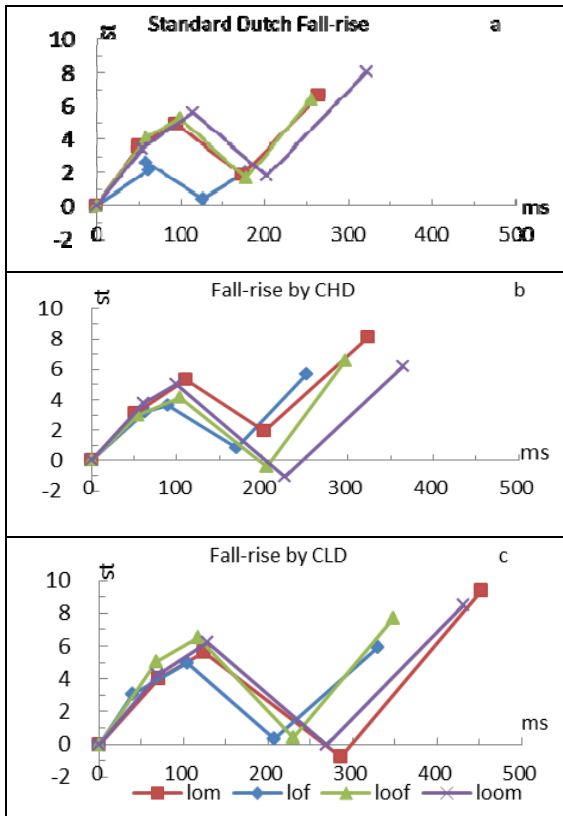


Figure 4: Fall-rises by native speakers (NSD) and Chinese speakers with higher (CHD) and lower (CLD) proficiency.

In the statistical analyses, we excluded the subjects who could not produce any one of the four target words correctly in the three contours. To get other missing data points, we selected only the subjects that had full data for the same variable in the four words and computed the means of the variable across the four words for each subject (row means) and the means of the same word across different subjects (column means). To fill in an empty cell, we took its corresponding column mean in the full data and added it to the mean difference between row mean of a speaker with missing data and the corresponding column means in the full data. We repeated this procedure for all other empty cells of the same variable, for all other variables in the dataset and for all datasets (Rise, Fall, Rise-fall), separately for native and Chinese speakers of Dutch.

3.2. Peak alignments in the Fall, Rise and Fall-rise

Peak alignment, a measure indicating the relative or absolute location of an f0 peak relative to the onset of a syllable or rime, depends on segmental phonetic and prosodic factors [1, 13]. We defined peak alignment as the peak position relative to the beginning of the rime and expressed it as the percentage of the total vowel duration (henceforth ‘peak delay’). Generally, the peak delay in Falls was greater for the three groups of speakers as the sonorant rime was longer. The differences in peak delay between the three groups are not significant. Although the percentages expressing to what extent the peak runs into the vowel are higher for NSDs than for CHDs, and those for CHDs are higher than those by CLDs, the differences between groups are not significant. Neither was a significant difference found when peak alignment was expressed as the absolute distance from the rime beginning. In the Rises, no significant peak delay differences were found, but we obtained marginally significant results when peak alignments were expressed in absolute terms. This result is due to the fact that rises tend to end where the syllable ends. Bearing this in mind, peak delays by LCD were larger than those by CHDs and those by CHDs were larger than those by NSD [$F(2, 38) = 3.2, p = .05, \eta^2 = .15$], with CLDs having longer peak delays ($p = .055$). For the Fall-rises, no significant group differences were found, either for our relative peak delay or for the absolute distances from rime beginnings.

3.3. Excursions of Fall, Rise and Fall-rise

A stable f0 excursion and an increased rate of f0 change on shorter words is interpreted as compression, while decreasing f0 excursions and a stable rate of f0 change is evidence of truncation.

The Fall. The order of f0 excursion ($H1-L2$) of the four target words by NSDs was $Lom_{H1-L2} > Loom_{H1-L2} > Loof_{H1-L2} > Lof_{H1-L2}$, and their rate of f0 change (RC) was ordered $Loof_{RC} > Lof_{RC} > Lom_{RC} > Loom_{RC}$, as shown in Figure 2. That is, the native speakers speeded up the rate of f0 change in shorter sonorant rimes, which means that they applied a compression strategy. In the utterances by CHDs, f0 excursions and rate of f0 change were ordered $Loof_{H1-L2} > Lom_{H1-L2} > Loom_{H1-L2} > Lof_{H1-L2}$ and $Lof_{RC} > Loof_{RC} > Lom_{RC} > Loom_{RC}$; in the CLDs’ utterances, this order was $Loom_{H1-L2} > Lom_{H1-L2} > Loof_{H1-L2} > Lof_{H1-L2}$ and $Lof_{RC} > Loof_{RC} > Loom_{RC} > Lom_{RC}$. For the Chinese subjects together the order for f0 duration was $Lof < Loof < Lom < Loom$. That is, like NSDs, CHDs compressed Falls. CLDs compressed *Loof* as much as *Lof*. No significant group differences were found in the f0 excursion and rate of f0 change. This means the pitch span differences between the groups were not significant.

The Rise. The f0 excursions for the NSD were $Lom_{L1-H1} > Loof_{L1-H1} > Loom_{L1-H1} > Lof_{L1-H1}$ and rates of change were ordered $Lof_{RC} > Loof_{RC} > Lom_{RC} > Loom_{RC}$. That is, on the basis of rate of change differences, native speakers compressed the Rise as sonorant portions got shorter. In the utterances by the CHD and CLD subjects, we found the orders $Lom_{L1-H1} > Loom_{L1-H1} > Loof_{L1-H1} > Lof_{L1-H1}$ and $Lof_{RC} > Loof_{RC} > Lom_{RC} > Loom_{RC}$ for CHD, while the CLDs had $Loom_{L1-H1} > Lom_{L1-H1} > Lof_{L1-H1} > Loof_{L1-H1}$ and $Lof_{RC} > Loom_{RC} > Loof_{RC} > Lom_{RC}$. Clearly, the performance by CHDs is quite similar to that of the native speakers, while the CLDs showed erratic orders when seen from the perspective of expected sonorant rime duration. No significant differences were found in f0 excursion between the three groups [$F(2, 38) = 1.1$, ins.], meaning that pitch range differences were not significantly different. However, the differences in rate of f0 change were significant [$F(2, 38) = 3.8$, $p < .05$, $\eta^2 = .168$]. Multiple comparisons showed that only the difference between NSD and CLD were significant ($p < .05$). More specifically, NSDs and CHDs increased the rate of f0 change when the sonorant rimes became shorter in a way that CLDs failed to follow. The CLDs had inadequate control of sonorant rime (see 3.1).

The Fall-rise. In the first peak of the Fall-rise (Figure 4), all three groups had longer f0 excursions in the target words with longer sonorant rimes, but only NSD and CHD had longer f0 excursions in *Loof* than in *Lom*. Differences in f0 excursion were significant between the three groups [$F(2, 38) = 6.4$, $p < .05$, $\eta^2 = .251$]. Values for LCD were the longest and those for NSD the shortest. Only the differences between CLD and NSD were significant. The differences in the rate of f0 change for the three groups were not significant

In the rising part of the contour, the differences in f0 excursions between the three groups were significant [$F(2, 38) = 6.4$, $p < .05$, $\eta^2 = .250$]. Multiple comparisons showed that only the difference between LCD and NSD was significant. As for the rates of f0 change in the second part of the contour by the three groups, the differences between the three groups were insignificant.

In both parts of the contour, the f0 excursions by NSD were significantly shorter than in the case of CLD, but no significant differences were found between NSD and CHD or between CLD and CHD. An inspection of Figure 4 reveals the reasons for these results. The NSD reduced the first peak and raised the following valley of the shortest sonorant rime, and allowed the final rise to end higher in the longest sonorant rime, but kept the fall-rise contours in all three longer sonorant rimes more or less intact. That is, neither CHD nor CLD reproduced the behaviour of NSD. In particular, neither the ordering of the two peaks nor that of the valleys in CHD follow the pattern of NSD, who neatly divided the four sonorant rimes into (a) shortest (*Lof*), medium (*Lom* and *Loof*) and longest (*Loof*), for the alignment and height of the first peak (later and higher when longer), the depth of the valley (lower when longer), and the height of the rise end (later when longer). No significant pitch span differences were found. Like the f0 of H1 in Fall (relative to L2) and Rise (relative to L1), differences in f0 between either H1 and H2 and L1 were insignificant.

4. Conclusion

We have presented evidence that, unlike less proficient Chinese learners of Dutch, more advanced learners produce Rises and Falls much the way native speakers do. As sonorant rimes are shorter, these contours are compressed by native

speakers and by the more proficient learners in our group of Mandarin Chinese subjects. The detailed patterning of these adjustments is not exactly the same, of course. In particular, in the case of the Fall, NSD end the fall for *Loom* earlier than in the longer *Lom*. In these data, the striking findings are that the lower proficiency group does not follow the pattern of NSD in any respect except that *Lom*, *Loom* have later falls than *Lof*, *Loof*. This suggests that, to them, there is no distinction between the long and short vowels of Dutch. Secondly, neither subgroup managed to reproduce the behaviour of NSD in the realization of the Fall-rise, which reflects the fact this pattern does not occur in their L1. While Tone 3 has a general falling-rising contour, in terms of Dutch phonology it is equivalent to a low rise, since the fall begins outside the rime instead of well within the rime (see Fig. 4). For this reason, the Fall-rise is more appropriately labeled a 'Rise-fall-rise'. We found no pitch span differences between the three groups. This contradicts [2], who showed that the pitch span by Chinese learners of English was narrower than that by native speakers. We provisionally attribute this difference to the fact that the pitch span of Dutch is narrower than that of English [e.g. 14 and references therein].

References

- [1] Caspers, J., Heuven, V. J. van (1993). Effects of time pressure on the phonetic realization of the Dutch accent lending pitch rise and fall. *Phonetica*, 50, 161–171.
- [2] Chen, H. (2008). On Chinese EFL learner's English intonation patterns. Shanghai: Shanghai Foreign Language Education Press.
- [3] Bannert, R., Bredvad, A. (1975). Temporal organization of Swedish tonal accent: the effect of vowel duration. Working papers 10, Phonetics laboratory, Department of General Linguistics, Lund University, Sweden.
- [4] Erikson, Y., Weenink, M. (1972). Fundamental frequency correlates of the grave word accent in Swedish: the effect of vowel duration. *Speech Transmission Laboratory, Quarterly Papers and Status Report 2-3* (pp. 53–60), KTH, Sweden.
- [5] Grabe, E. (1998). Pitch accent realization in English and German. *Journal of Phonetics*, 26, 129–143.
- [6] Grosser, W. (1997). On the acquisition of tonal and accentual features of English by Austrian learners. In A. James, J. Leather (eds.), *Second Language Speech: Structure and Process* (pp. 211–228). Berlin: Mouton de Gruyter.
- [7] Gussenhoven, C. (2005). Transcription of Dutch intonation. In S.-A. Jun (ed.), *Prosodic typology: the phonology of intonation and phrasing* (118–145). Oxford: Oxford University Press.
- [8] Gut, U. (2009). *Non-native speech*. Frankfurt: Peter Lang.
- [9] Hanssen, J., Peters, J., Gussenhoven, C. (2007). Phrase-final pitch accommodation effects in Dutch. *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, 1077–1080*.
- [10] Ladd, D. R. (2008). *Intonational phonology*. Cambridge: Cambridge University Press.
- [11] Peng Shu-hui, Chan, M. K. M., Tseng Chiu-yu, Huang Tsan, Lee Ok Joo, Beckman, M. E.. (2005). Towards a Pan-Mandarin system for prosodic transcription. In: Jun, S.-A. (ed.), *Prosodic typology: the phonology of intonation and phrasing* (pp. 230–270). Oxford: Oxford University Press.
- [12] Warren, M. (2004). A corpus-driven analysis of the use of intonation to assert dominance and control. In U. Connor, T. Upton (eds.) *Applied Corpus Linguistics: A Multidimensional Perspective* (pp. 21–33). Amsterdam: Rodopi.
- [13] Xu, Y. (2001). Fundamental frequency peak delay in Mandarin. *Phonetica*, 58, 26–52.
- [14] Willems, N., Collier, R., Hart, J. 't. (1988). A synthesis scheme for British English intonation. *Journal of the Acoustical Society of America*, 84, 1250–1261.