

PHONETIC IMPLEMENTATION MUST BE LEARNT: NATIVE VERSUS CHINESE REALIZATION OF FOCUS ACCENT IN DUTCH

Xuliang He^{a,b}, Judith Hanssen^b, Vincent J. van Heuven^c & Carlos Gussenhoven^{b,d}

^aSchool of Foreign Studies, Nantong University, China;

^bCentre for Language Studies, Radboud University Nijmegen, the Netherlands;

^cPhonetics Laboratory, Leiden University Centre for Linguistics, the Netherlands;

^dQueen Mary University of London, UK

x.he@let.ru.nl; j.hanssen@let.ru.nl; c.gussenhoven@let.ru.nl;

v.j.j.p.van.heuven@hum.leidenuniv.nl

ABSTRACT

The strategies adopted by native speakers of Dutch and Chinese speakers with different levels of proficiency in Dutch in the realization of degrees of focus enhancement were significantly different on six variables. The participants with higher proficiency did not outperform the learners with lower proficiency. The Chinese subjects appear to abstain from phonetic enhancement rather than transferring L1 patterns of realization.

Keywords: enhancement of focus, Chinese L2 Dutch

1. INTRODUCTION

Focus constituents vary in size. They may be ‘broad’, as in the Dutch (1a), or ‘narrow’, as in (1b,c). In both cases, the focus constituents provide the information requested in the precursor sentence [9]. Focus has also been divided up into different types [5]. The focus constituent *Momberen* in (1b) has the same size as that in (1c), but has been distinguished as ‘informational’ and ‘corrective’, respectively. The *Momberen* in (1b) is supplied to fill an informationally empty slot, but *Momberen* in (1c) is a correction of existing information, *Zaltbommel*.

(1a) *Wat zijn je plannen voor morgen?*

‘What are your plans for tomorrow?’

Ik [zou wel naar Momberen] willen fietsen

‘I would like to cycle to Momberen’

(1b) *Waar zou Karel je heen willen brengen?*

‘Where would Karel like to take you?’

Hij zou me naar [Momberen] willen brengen.

‘He would like to take me to Momberen’

(1c) *Wilde je moeder je naar Zaltbommel sturen?*

‘Did your mother want you to go to Zaltbommel?’

Nee ze had me naar [Momberen] willen sturen.

‘No, she wanted to send me to Momberen’

Speakers of different languages adopt phonetically and phonologically different strategies to express the different focus conditions as shown in (1). These strategies include the use of different pitch accents, differences in prosodic phrasing, differences in pitch range and duration differences (e.g. [12, 14]). Standard Dutch has been reported to display focus-related differences in the duration of onsets and codas of the nuclear accented syllables, in f₀ scaling and in the timing and slope of the pitch fall in H*L L% contours [7].

The aim of this contribution is to investigate the extent to which such – largely subtle – phonetic adjustments are picked up by foreign learners. Many prosodic differences between native speech and L2 speech can be characterized as phonetic where they are due to different realizations of comparable phonological categories [6, 10]. Northern Chinese L2 speakers of Dutch find it relatively easy to judge the correctness of intonation contours of deaccented postfocal speech [8], which is due to the fact that Beijing Mandarin has comparable postfocal Pitch Range Compression (PFC) [13]. Also, Beijing Mandarin employs longer syllable durations on focused words as well as an expanded f₀ range [1, 3, 14]. Broadly, therefore, the articulatory behaviour of northern Chinese speakers of Dutch may be expected to come close to that of native speakers, when realizing the target sentences in each of the three focus conditions with H*L pitch accents. However, it is not evident that they will display the same variation across the three focus conditions phonetically. Our research question, therefore, is whether native L1 and northern Chinese L2 speakers of Dutch realize the three focus conditions in (1a,b,c) in comparable ways, and if not, whether more advanced learners come closer to the native behaviour than less advanced learners.

2. EXPERIMENT

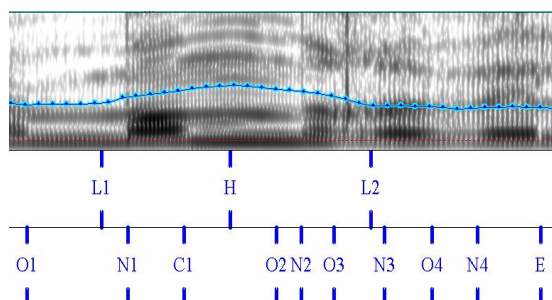
Three sets of short declarative sentences were used to elicit broad, narrow, and corrective focus. Each focus set had four target words, the pseudo placenames *Manderen*, *Momberen*, *Memberen* and *Munderen*. Examples of contexts and target sentences for *Manderen* were given in (1). Before the recording, subjects heard two example dialogues which were pre-recorded by native speakers of Dutch. These dialogues were subsequently presented to them in written form, and only if subjects correctly accented the target words in these practice dialogues did we proceed with the recording of the experimental dialogues. The 12 dialogues were presented in random order. Participants were allowed to repeat sentences whenever they liked. The subject's voice production was recorded with a Zoom H4 Handy Digital Recorder.

2.1. Subjects

Twenty Chinese speakers (3 male) of Dutch, aged between 17 and 53, took part in the 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 [8]. The native speakers (NSD) in the corpus were 8 males and 13 females, aged from 14 to 49.

2.2. Measurements and variables

Figure 1: f0 minima and maxima were labeled on the first tier and segmental boundaries of the test word 'Manderen' (O1-O4) and its post word 'willen' (O4-E) on the second.



As Figure 1 shows, f0 minima and maxima within the target words were established, automatically for the f0 maxima and manually for the f0 minima at the beginning of the contour (L1) and at the end of the Fall (L2) at the point in the pitch signal where a sudden change in slope is apparent. The segmental boundaries of the test

words were labeled at the negative-to-positive zero-crossings based on auditory information and visual inspection of the spectrogram and the waveform.

In view of our aim to detect differences in the fine detail of the phonetic implementation in the three focus conditions as a function of participant group, 23 variables were defined. For example, the durations of the onset (O1-V1), coda (C1-O2), the rime (N1-O2) and the Fall duration (H-L2), the f0 excursion of nuclear Fall (H-O2), and the timing of the peak relative to the beginning to the vowel (peak delay) as a proportion of the rime duration.

3. RESULTS

3.1. Results for the correct readings

Despite the brief training session, participants produced 52 (21.7%) utterances that could not be used due to phonological errors of various kinds. Of these, 56.5% were due to incorrect stress placement on the target word, 35.5% to deaccentuation of the target word or incorrect accentuation of the post-target word. Further, two participants made incorrect pitch accent choices and three made other pronunciation errors. One CHD speaker pronounced all the target words in the corrective and narrow focus conditions incorrectly. Three CLD speakers were incapable of pronouncing the target words in one or more focus conditions. Data from the participants were discarded. Other missing data were replaced with the means in the same focus condition.

3.2 Comparisons of focus effects

We ran Repeated Measures ANOVAs for all variables with focus condition (3 levels) as a within-subject factor and groups as a between-subjects variable. Only six of 23 variables that were found to be significantly affected. Huynh-Feldt corrected p -values are reported whenever the assumption of sphericity was violated. Only significant interactions between focus condition and group for these six variables are analyzed.

Onset duration. While the focus condition did not yield significantly different results within groups, the interaction between groups and focus condition is significant [$F(4,68)=2.6$, $p<.05$, $\eta^2=.10$]. From broad to narrow and corrective focus, both L2 groups decreased the onset duration, but L1 speakers increased it by 7 ms going from broad to narrow and corrective. There was a main effect for group [$F(2,34)=4.6$, $p<.05$,

$\eta^2=.20$], with Bonferroni multiple comparisons showing significantly longer onset duration by CLDs than by CHDs.

Rime duration. A significant interaction between group and focus was found [$F(4,68)=3.4$, $p<.05$, $\eta^2=.17$]. The rime duration by CHDs decreases from broad to narrow to corrective focus, a pattern found in neither of the other groups. NSDs have longer rimes in narrow focus, but shorter rimes in corrective focus, than in neutral focus; for CLDs, narrow focus yielded longer rimes than the other two conditions.

Coda duration. A main effect of focus was found [$F(2,68)=3.3$, $p<.05$, $\eta^2=.10$] as well as an interaction between group and focus [$F(4,68)=2.8$, $p<.05$, $\eta^2=.10$]. Between subjects, a main effect of group [$F(2,34)=8.8$, $p<.01$, $\eta^2=.30$] was largely due to the fact that coda durations in the NSDs are significantly longer than those by CHDs and CLDs. No significant differences were found between the latter groups.

Nuclear fall excursion. A significant interaction was found between group and focus [$F(4,68)=3.8$, $p<.05$, $\eta^2=.20$]. The nuclear fall excursions by NSDs increase from broad to narrow to corrective focus, but those by CHDs decrease. For CLDs, no regular pattern was found. The differences in nuclear fall excursion between groups are significant [$F(2,68)=4.7$, $p<.05$, $\eta^2=.20$]. Multiple comparisons showed that excursions by NSDs were significantly greater than those by CLDs, but those by CHDs did not significantly differ from those of the other groups.

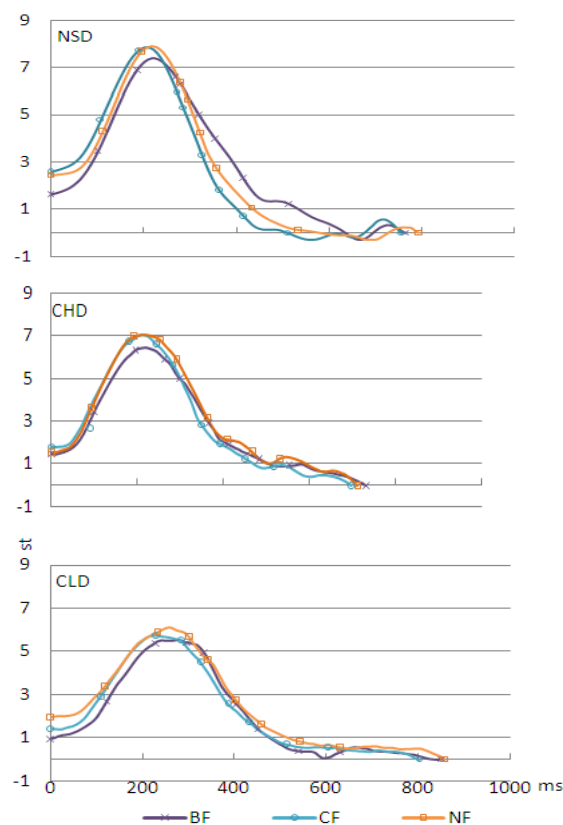
Peak delay. A significant interaction between group and focus was found [$F(4,64.7)=4.6$, $p<.05$, $\eta^2=.20$]. Relative peak delay decreases from broad to narrow to corrective for the NSDs. For CLDs, it is largest under broad focus and smallest under narrow focus, while for CHDs the opposite is true. Overall differences between groups are also significant [$F(2,34)=4.0$, $p<.05$, $\eta^2=.20$], with CHDs' peak delays being later than those by NSDs.

Fall duration. A significant interaction between group and focus was found [$F(2.8,47.8)=3.8$, $p<.05$, $\eta^2=.20$]. The Fall durations for NSDs decrease from broad to narrow to corrective focus but increase for the CLDs. CHDs are similar to NSDs, but they have a longer Fall duration in corrective focus than narrow focus. The effect of group is significant [$F(2,34)=3.8$, $p<.05$, $\eta^2=.20$], but only CHD and NSD differ from each other.

Figure 2 shows the average pitch contours in the three focus conditions for NSDs, CHDs and CLDs

separately. Here the information for one hundred data points from the onset of the target word to the end of the post-nuclear word was averaged over subjects.

Figure 2: Averaged contours under three focus conditions (Broad focus: BF, Corrective focus: CF, Narrow focus: NF) by native speakers of Standard Dutch (NSD), Chinese speakers of Dutch with higher proficiency (CHD) and Chinese speakers of Dutch with lower proficiency in Dutch (CLD). The positions of the crosses, circles and squares are the segmental boundaries in the target word and the post-nuclear word. F0 (semitones) is expressed relative to the end of the post-nuclear word.



4. DISCUSSION

Overall these results indicate that the phonetic detail in the realization of H*L falling contours in IP-internal syllables is specific to Standard Dutch. Chinese speakers of Dutch perform pitch falls that resemble the Dutch pitch falls both in shape and approximate timing, but neither the detailed phonetics of these falls nor the systematic variation across different focus conditions match the phonetic patterning of the native speakers. Summarizing the findings for the native speakers, we find that, going from broad informational to narrow (informational or corrective) focus, NSDs increase the onset duration. Tendencies to *lengthen*

onset consonants under stress have been widely noted (e.g. [4] for English). Second, the *excursion of the Fall* increases along the same dimension, a commonly reported effect of emphasis and focus. Third, *peak delay*, the relative timing of the peak within the rime, decreases along the same dimension. Early peak placements under enhancing focus conditions have been reported for other languages [11]. Even when peak delay is expressed in absolute distance from the rime beginning rather than in terms of a percentage of the rime duration, the regularity holds for the NSDs (from broad via narrow to contrastive the absolute values are 118, 104 and 99 ms). Fourth, again going from broad via narrow to the corrective focus condition, NSDs decrease the duration of the Fall, so that it is steeper as the focus is more emphatic. Partly similar data were interpreted by [7] as hyper-articulation of the H*L pitch accent. Because the rising movement towards the peak remains unaffected, [7] took this finding as evidence for the analysis of the pitch peak as due to a H*L pitch accent as opposed to a LH* pitch accent. In the data presented here, the regularity is stronger, with a gradient decrease in Fall duration across the three focus conditions. The NSDs, therefore, produced higher, earlier and faster Falls as the communicative situation demanded greater articulatory precision. One phonetic measure failed to show the same gradient pattern. The NSDs produced the longest coda and rime durations in the intermediate narrow informational condition and the shortest in the corrective narrow focus condition. This shows that rime lengthening is not necessarily a concomitant of hyper-articulation. Rather, we suggest that it can be, as shown by the difference between the broad and narrow informational focus conditions, but that the effort to hyper-articulate the fall by making it steeper led to a shortening not just of the falling pitch movement but also of the rime on which it was executed. The fine phonetic detail in the behavior of the NSDs thus constitutes a coherent set of measures to increase articulatory precision.

By comparison, the Chinese speakers of Dutch display irregular behavior. Even though there are many significant differences for the measures we have chosen, these do not obviously add up to a coherent pattern. Perhaps it is best characterized as uncertain. Significantly, [2] found that speakers of Standard Chinese do display regular hyper-articulation responses in a similar task. The fact that such responses are in part a matter of choice

may mean that when speaking a foreign language, speakers will at first not commit themselves to any specific response. The uncertain patterns may conceivably contribute to their foreign accent. Unlike what was found by [8], the proficient learners did not always do better than the less proficient ones in this production test. The phonetic realization of focus conditions should evidently be learnt.

5. REFERENCES

- [1] Chen, Y. 2006. Durational adjustment under corrective focus in Standard Chinese. *Journal of Phonetics* 34, 176-201.
- [2] Chen Y., Gussenhoven, C. 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36, 724-746.
- [3] Greif, M. 2010. Contrastive focus in Mandarin Chinese. *Proc. Speech Prosody*. Chicago.
- [4] Gussenhoven, C. 1999. Discreteness and gradience in intonational contrasts. *Language and Speech* 42, 283-305.
- [5] Gussenhoven, C. 2007. Types of focus in English. In Lee, C. M., Gordon, M., Büring, D. (eds.), *Topic and Cross-Linguistic Perspectives on Meaning and Intonation*. Dordrecht: Springer, 83-100.
- [6] Gut, U. 2009. *Non-native speech*. Frankfurt: Peter Lang.
- [7] Hanssen, J., Peters, J., Gussenhoven, C. 2008. Prosodic effects of focus in Dutch declaratives. *Proc. Speech Prosody*. Campiñas, Brazil, 609-612.
- [8] He, Xuliang, van Heuven, V., Gussenhoven, C. 2010. Choosing the optimal pitch accent location in Dutch by Chinese learners and native listeners. *Proc. 6th New Sounds* Poznań, Poland.
- [9] Ladd, D.R. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press.
- [10] O'Brien, M., Gut, U. 2010. Phonological and phonetic realization of different types of focus in L2 speech. *Proc. New Sounds* Poznań, Poland.
- [11] Smiljanić, R., Hualde, J.I. 2000. Lexical and pragmatic functions of tonal alignment in two varieties of Serbo-Croatian. *Chicago Linguistic Society* 36, Vol. 1, 469-482.
- [12] Ueyama, Motoko, Jun, S.-A. 1998. Focus realization in Japanese English and Korean English intonation. *Japanese and Korean Linguistics* 7, 629-645.
- [13] Wu, W.L., Xu, Y. 2010. Prosodic focus in Hong Kong Cantonese without post-focus compression. *Proc. Speech Prosody* Chicago.
- [14] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27, 55-105.