THE REALISATION OF NUCLEUS PLACEMENT IN SECOND LANGUAGE INTONATION

Ineke Mennen
University of Edinburgh & Queen Margaret University College, Edinburgh, UK

ABSTRACT

This paper investigates the realisation of nucleus placement in the production of Greek yes/no questions by non-native (Dutch) advanced speakers of (Modern) Greek. The yes/no questions which are investigated can have two possible nucleus locations: the nucleus can appear on the first or the last of two content words. It will be shown that half of the subjects of this study do not distinguish between the two in production, and consistently produce either just one or the other type. Experimental evidence will be given that the contour which does not have a counterpart in the first language (L1), is more accurately produced than the one which is more similar to an L1 contour. However, both contours are phonetically inaccurately produced when compared to the native norm. It will be shown that this is possibly caused by a phonological misinterpretation of the contour's final rise-fall.

1. INTRODUCTION

Greek yes/no questions (YNQs) can autosegmentally be described as L* H L%, that is as consisting of low F0 on the nuclear syllable (L*), followed by a H L% boundary sequence, which is manifested by a rise-fall towards the end of the utterance [1]. Although the peak (H) of this final boundary sequence always occurs near the end of the sentence, its exact location depends on the location of the nuclear accent. If the nuclear accent occurs on the last word, the H occurs on the utterance-final syllable (regardless of whether this syllable is stressed or not). If, however, the nuclear accent is on an earlier word, the H occurs on the stressed syllable of the last word. In other words, the exact realisation of the L* H L% sequence in Greek YNQs depends on (i) the location of the nucleus, and (ii) the location of the lexical stress of the utterance-final word. An example of a nucleus-final (NF) question, is given in Figure 1. In Figure 2 an example of a nucleus non-final (NNF) YNQ is shown.

The most obvious feature of Dutch YNQ intonation is its rise at the end of the utterance, present in the majority of YNQs [2]. Dutch YNQs can be described autosegmentally as the sequence H* L H% or L* H% [e.g. 3]. In other words, the nuclear accent in Dutch YNQs can either be L* or H*, and the final boundary tone is usually H%. If the nuclear accent is H*, F0 is high on the nuclear syllable, and begins to drop in that same syllable [4]. If it is low, the F0 drop starts earlier, just before the accented syllable [3]. In both cases it remains low until the final syllable where it usually starts rising again. Figure 3 illustrates the Dutch H* L H% yes/no question. An example of the L* H% yes/no question is shown in Figure 4.

It is obvious from the above description and the Figures 1 to 4, that there are cross-linguistic differences between YNQ intonation in Dutch and Greek. One such difference is the realisation of rise-falls. Although in both languages a rise-fall can occur in YNQ intonation, in Dutch it can never occur on an unstressed syllable, as is the case in Greek NF YNQs. Furthermore, it is likely that Dutch listeners, when listening to Greek YNQs, will perceive the syllable bearing the final rise-fall as the most prominent. For Greek speakers, however, it is the low-pitched lexically stressed syllable which is perceived as the most prominent in the utterance [1].

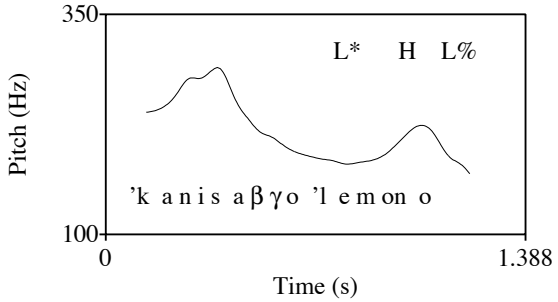

Figure 1. An example of a Greek NF contour [ˈkanis aβɤoˈlemono] 'Did you make EGG AND LEMON SAUCE?', where focus is on the last content word (the underlined word in the transcription, the word in bold in the gloss).

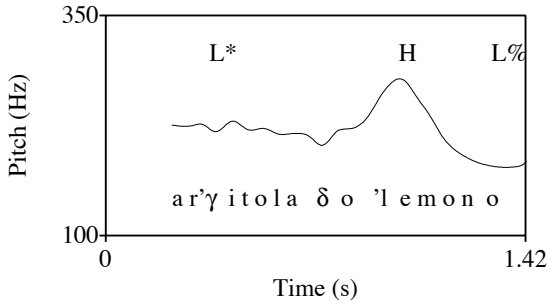

Figure 2. An example of a Greek NNF YNQ contour [arˈɤi to laðoˈlemono] 'Does the OIL AND LEMON SAUCE take long?', where focus is on the first content word (the underlined word in the transcription, the word in bold in the gloss).

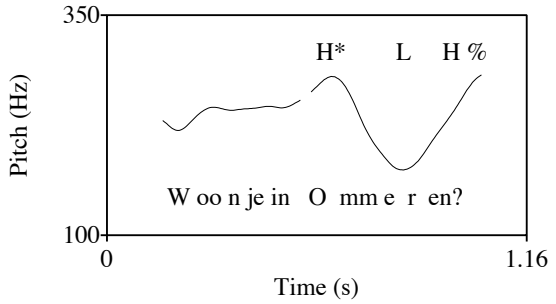

Figure 3. A Dutch H* L H% yes/no question contour of the sentence [ˈʊoŋə ɪn ˈomərə]'Do you live in Ommeren? F0 is high on the nuclear syllable and starts to drop in that same syllable.

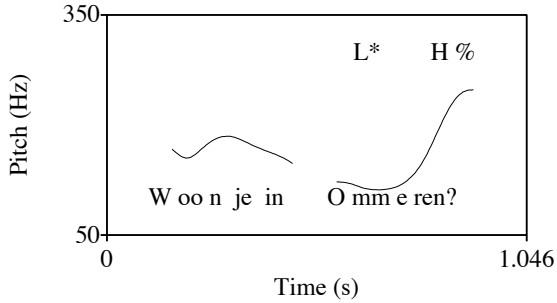

Figure 4. A Dutch L* H% yes/no question contour of the sentence [ˈʊoŋə ɪn ˈomərə] 'Do you live in Ommeren? F0 is already low at the nuclear syllable.

1.1. Flege's Speech Learning Model (SLM)

This model was developed to account for segmental aspects of language learning, and is based on 'acoustic-phonetic' similarity between L1 and L2 sounds [5]. It has been shown elsewhere that this model can also help account for prosodic aspects of L2 speech [6]. According to the SLM, the degree of perceived similarity between L1 and L2 sounds will determine whether new categories for L2 sounds can be established. When L1 and L2 sounds are acoustically 'similar' (but not identical), both sounds will be classified into a single category, which will result in accented production of the L2 sound. When the L2 sound is noticeably different from any L1 sound, i.e. when it is 'new', it should be possible to establish a new category for this sound. Production and perception of this sound should be fairly unproblematic, although not necessarily identical to that of native speakers [5]. If applied to intonation, the SLM would predict that the L2 learner would have more difficulty with the production of a 'similar' than a 'new' contour. In the case of Dutch learners acquiring Greek YNQs, the distinction between 'new' and 'similar' is not that straightforward. However, it may be that one of the different focus readings in the Greek YNQs can be seen as 'new' and the other as 'similar'. A 'similar' contour is the one observed when the first content word is focused, i.e. the NNF condition. In this case the final rise-fall occurs on the lexically stressed syllable of the utterance-final word. As in Dutch a rise-fall always associates with a lexically stressed syllable, this pattern is similar to Dutch. However, in Dutch it would usually occur in statements, rather than yes/no questions. A rise-fall on an unstressed syllable, on the other hand, is not possible in Dutch. Therefore, the NF pattern should be seen as 'new'. As a consequence Dutch learners of Greek should succeed better in producing the new NF contour than the similar NNF contour. The following experiment was conducted in order to test this hypothesis.

2. METHOD

2.1. Materials

The materials consisted of two sets of 60 short Greek YNQs containing two content words, each set reflecting a different focus placement. This focus placement resulted in the experimental condition 'nucleus location'. One set consisted of nucleus-non-final (NNF) test sentences, where focus was expected to be on the first content word. The other set consisted of nucleus-final sentences (NF), in which focus was expected on the last content word. To achieve the desired nucleus location, each test sentence was presented on a card in a short dialogue. In order to obtain smooth F0 contours, the test words consisted mostly of sonorants.

2.2. Subjects and procedure

Two groups of subjects were used:

- Group DG: six non-native (Dutch) speakers of Greek, three males (DG1, DG3, DG5), and three females (DG2, DG4, DG6). They were all very advanced (near-native) speakers of Greek, taught Greek at University level, and had between 6 and 35 years of experience in the L2. All speakers had started learning Greek in adulthood. The speakers differed considerably in age, with one speaker in her twenties, one in his fifties and the others in their thirties and forties. For reasons of privacy the age and amount of experience with the L2 of each individual speaker will not be revealed here.
- Group G: five native speakers of Greek, three females (G3, G4, and G7) and two males (G6 and G10). All speakers were recruited from the Edinburgh student population. All speakers of both groups were reasonably competent in English as well. The speakers were all in their twenties and thirties.

For the recording the speakers read the dialogues from cards which were interspersed with materials for other experiments (not reported here). The materials were recorded on DAT tape and digitised at a 16kHz sampling rate with appropriate low-pass prefiltering using a SunSPARC workstation with ESPS Waves+ speech analysis facilities.

2.3. Measurements and analysis

For each speaker the same 45 sentences for each nucleus location condition were selected for further measurement, yielding a total

of 90 sentences for each speaker. Sentences were not selected when they were disfluent, or when they were produced with another intonation contour (e.g. a statement instead of a YNQ) than the one intended. Visual inspection of the pitch tracks suggested that the main difference between the two groups lay in the *scaling* and in the *alignment* (timing) of the peak of the final rise-fall with segments. Other less striking differences will not be discussed here, due to space limitations. The measurements are defined as:

- Scaling of H: the highest *F*0 point of the utterance-final rise-fall
- Alignment of H: the distance (in ms) between the offset of the stressed vowel of the final content word and the H

It was hypothesised that Group DG would experience difficulty with the final rise-fall of the Greek YNQs. If, for example, differences were found in the scaling of the final H between native and non-native speakers, it was crucial to be able to decide whether these differences could be attributed to real group differences, or whether they were caused by individual speakers' differences in pitch range. For this reason, it was thought necessary to abstract away from differences between speakers. Therefore, F0 values were expressed on a speaker-specific scale which was obtained by assigning a value of 100 to the top and a value of 0 the bottom of the speakers overall F0 range. The scaling measurements were then expressed on this speaker-specific scale. For a similar approach see [7, 8].

3. RESULTS

All results here are based on mixed analyses of variance (within and between items). There are three factors: group, speaker, and nucleus location. The factor group is within-items and has two levels (DG, G). The factor speaker is within-items, has eleven levels, and is nested within the factor group. The factor nucleus location is between-items and has two levels (NF, NNF).

Nucl. location	DG1	DG2	DG3	DG4	DG5	DG6	G
NF:							
Mean	103	-73	121	80	92	96	58
S.E.	21	12	19	17	21	23	22
NNF:							
Mean	-44	-90	118	-75	85	-65	-89
S.E.	8	7	21	7	21	9	10

Table I. Means in (ms) of the alignment of H (the distance between the offset of the stressed vowel of the final content word and the H). Means and standard errors are presented for all speakers of Group DG in both nucleus locations (NNF and NF), along with the pooled means for the speakers of Group G.

3.1. Nucleus location

Although no instructions were given about where to place the nuclear accent, in the majority of items (99%) speakers of Group G placed the accent where it was intended in the design of the materials. For speakers of Group DG, agreement appeared not to be that high, when judged auditorily. Therefore, before group analyses were carried out it was established on the basis of the alignment measurement (the interval between the offset of the stressed vowel of the final content word to the H) whether the speakers of Group DG had produced the intended nucleus placement and whether they produced a distinction between the NF and NNF condition. Table I shows mean alignment values for each of the speakers of Group DG and each nucleus condition, along with the overall means for group G. Separate analyses of variance (ANOVAs) for each of the speakers of Group DG showed that only half of the speakers distinguished between the two (for DG1 [$F(1,88) = 42.722, p < 0.0001$], for DG4 [$F(1,88) = 72.158$, $p < 0.0001$], and for DG6 [$F(1,88) = 41.382$, $p < 0.0001$]), i.e. the interval was negative in the NNF condition and positive in the NF condition, just as for Group G ([$F(1,88) = 36.170$, $p < 0.0001$],. For the other three speakers there was no effect of nucleus location. This indicates that these speakers consistently produce just one nucleus condition, that is speaker DG2 invariably produces all sentences with the nucleus on the first content word (NNF), whereas speakers DG3 and DG5 always produce NF YNQs.

For the comparison of NF with NNF conditions, means were calculated for each item, averaging over the speakers of each group, and were entered into ANOVAs. The group means for Group DG are based only on those speakers that distinguish between these two conditions, i.e. speakers DG1, DG4 and DG6, whereas the group means for Group G are based on all five speakers. The speakers of Group DG who do not distinguish between NF and NNF show a similar pattern than the other speakers of Group DG for the one condition they produce. Their means were entered into separate ANOVAs and compared to the group means of Group G.

3.2. Scaling of the H

The results of an ANOVA (group x nucleus location) show that there is a main effect of both group and nucleus placement. That is, the H is scaled higher in NNF than in NF condition [$F(1,88) = 190.466$, $p < 0.0001$], and it is higher for Group DG than for Group G [$F(1,88) = 1038.972$, $p < 0.0001$]. Furthermore, there is an interaction between the two factors, indicating that the H in the NNF condition is further removed from the means of Group G, than in the NF condition [$F(1,88) = 23.040$, $p < 0.0001$], as shown in Figure 5. This is exactly as expected from the hypothesis that the 'new' contour (i.e. the NF condition) should be easier to produce than the 'similar' NNF contour.

The results on the NF condition for the speakers DG3 and DG5 (who consistently produce the NF contours only) indicate that they show a similar pattern to the other speakers of Group DG. For both speakers the H is scaled significantly higher than for Group G (for speaker DG3 [$F(1,44) = 897.217, p < 0.0001$]; for DG5 [$F(1,44) = 72.160, p < 0.0001$].

Speaker DG2 (who consistently produces the NNF contour only) also shows a similar pattern to the other speakers of Group DG in the NNF condition. The H is scaled significantly higher than for Group G[$F(1,44) = 172.140, p < 0.0001$].

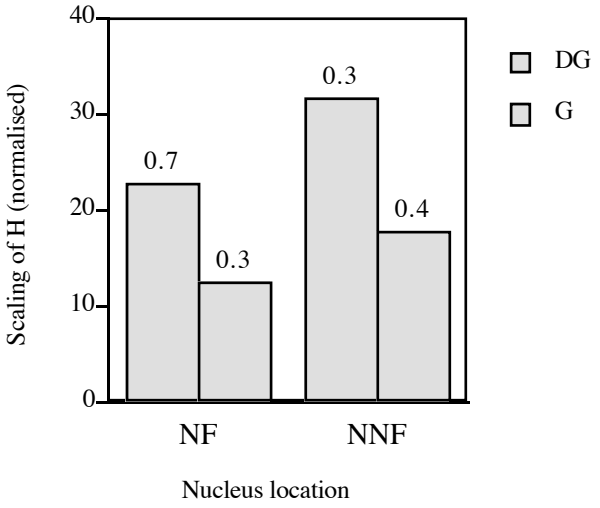

Figure 5. Scaling of the H for the two groups of speakers. The H is higher for Group DG than for Group G, especially in the NNF condition. The numbers above the bars represent the standard errors.

4. DISCUSSION AND CONCLUSION

Overall, the data show that there are differences in the production of Greek YNQs between native and non-native (Dutch) speakers of Greek. Specifically, it is shown that in their production of Greek YNQs only half of the subjects produce the two different nucleus locations (which reflect differences in focus) which are made by native speakers. Instead, they consistently produce just one of the two possible nucleus locations (NF or NNF). It is unclear whether this is due to the fact that they cannot produce a distinction, or whether they fail to recognise from the supplied dialogues which nucleus placement was required. In the latter case the problem may be pragmatic rather than phonological or phonetic. Even so, it is rather surprising that very advanced speakers of Greek fail to recognise a distinction which was recognised by 99% of the native speakers, especially after such long exposure to the L2.

For the speakers who produced a distinction between the two types of nucleus location, it is shown that the NF contour, which does not have a counterpart in the first language (L1), is more accurately produced than the NNF contour, which is more similar to the L1. This is just as predicted by the SLM.

Nevertheless, it is shown that the speakers of Group DG also failed to produce the 'new' contour accurately. This may be due to the fact that the final rise-fall in both cases of Greek YNQs, is in some sense 'new'. That is, in Greek this rise-fall occurs after the nuclear accent. In Dutch, a rise-fall can only occur on or before the nuclear accent. Thus, it is likely that the speakers of Group DG perceive the syllable bearing the final rise-fall as the nuclear syllable, and fail to recognise that this rise-fall is in fact a boundary sequence. It is therefore likely that it is this phonological misinterpretation which causes the differences in scaling between native and non-native speakers.

ACKNOWLEDGMENTS

The data for the native speakers reported here were collected for a project funded by the U.K. Economic and Social Research Council through grant no. R000 23 5614. This support is gratefully acknowledged. Furthermore, I would like to thank the speakers who participated in this study, most of all the non-native speakers, for allowing me to report on their L2 production data.

REFERENCES

[1] Arvaniti, A., Ladd, D.R. and Mennen, I. Quantitative F0 descriptions require reference to local targets: the intonation of Greek polar questions and emphatic statements (forthcoming).

[2] Haan, J., van Heuven, V.J., Pacilly, J.J.A. and van Bezooijen, R. 1997. An anatomy of Dutch question intonation. In De Hoop, H. and Coerts, J. (eds.), *Linguistics in the Netherlands*. Amsterdam: John Benjamins.

[3] Gussenhoven, C. 1988. Adequacy in intonation analysis: The case of Dutch. In Van der Hulst, H. and Smith, N. (eds.), *Autosegmental studies on pitch accent*. Dordrecht-Cinnaminson: Foris Publications.

[4] Collier, R. and 't Hart, J. 1981. *Cursus Nederlandse Intonatie.* Amsersfoort: Acco.

[5] Flege, J.E. 1995. Second language speech learning: Theory, findings and problems. In Strange, W. (ed.), *Speech perception and linguistic experience: Issues in cross-language speech research*. Timonium MD: York Press.

[6] Mennen, I. Second language acquisition of intonation: the case of peak alignment. In Gruber, M.C., Higgins, D., Olson, K. and Wysocki, T. (eds.), *Chicago Linguistic Society 34 II: The panels* (in press).

[7] Earle, M.A. 1975. An acoustic phonetic study of Northern Vietnamese tones, Speech Communication Research Laboratory, Santa Barbara.

[8] Rose, P. 1987. Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication* 6: 343-51.