Language-specific Effects of Pitch Range on the Perception of Universal Intonational Meaning

Aoju Chen, Toni Rietveld and Carlos Gussenhoven

Centre for Language Studies
University of Nijmegen, The Netherlands
aoju.chen@let.kun.nl; a.rietveld@let.kun.nl; c.gussenhoven@let.kun.nl

Abstract

Two groups of listeners, with Dutch and British English as their native language judged stimuli in Dutch and British English, respectively, on the scales CONFIDENT vs. NOT CONFIDENT and FRIENDLY vs. NOT FRIENDLY, two meanings derived from Ohala's universal Frequency Code. The stimuli, which were lexically equivalent, were varied in pitch contour and pitch range. In both languages, the perceived degree of confidence decreases and that of friendliness increases when the pitch range is raised, as predicted by the Frequency Code. However, at identical pitch ranges, British English is perceived as more confident and more friendly than Dutch. We argue that this difference in degree of the use of the Frequency Code is due to the difference in the standard pitch ranges of Dutch and British English.

1. Introduction

In a study on the communicative effects of rising and falling intonation in British English (henceforth BrE) and Dutch, Rietveld et al. [1] assessed a number of attributes of intonation contours H*LL% and L*HH%. It was found that in BrE, H*LL% was perceived as more negative, while in Dutch, L*HH% was perceived as more negative. In their experiment, the same pitch range was used for the BrE and the Dutch stimuli. Since BrE has wider pitch excursions than Dutch, the pitch range adopted in the experiment might have sounded normal to BrE listeners, but strikingly high to Dutch listeners. The authors therefore suggest that the difference in their results is caused by pitch range differences.

The present study aims to investigate the effects of pitch range on the perception of the universal meaning of intonation. The universal meaning at issue is the Frequency Code [2,3,4], which is based on the fact that smaller larynxes produce higher notes than larger ones. One of its manifestations is that high pitch sounds vulnerable and submissive, while low pitch sounds protective and dominant. Hence, high, and particularly high-ending utterances sound dependent, uncertain, appealing etc.; low and low-ending utterances sound powerful, assertive and authoritative. We know from other work [5] that the Frequency Code appears to be quite generally observed in languages. However, if we were to find that two languages use the Frequency Code differently, we would have found a language-specific meaning component. Suppose that listeners of two languages, Language A and Language B, are asked to judge a number of utterances in their native language, varying in average F0 values, with the help of a semantic attribute derived from the Frequency Code, like SUBMISSIVENESS. Assume for the sake of the argument that the two languages have identical mean pitch ranges, meaning that the difference between the lowest and the highest pitch ranges are equal in the two languages. We will use the term 'standard pitch range' for the mean pitch range of a language. Theoretically, the use of the Frequency Code in the two languages can differ at least in three ways:

a. Type 1 difference: In both languages, the degree of submissiveness increases when the pitch range is raised. However, the increase is stronger in Language A than in Language B, as illustrated in Figure 1.

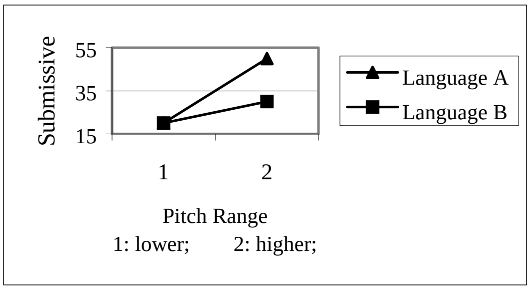

Figure 1. Type 1 difference

b. Type 2 difference: In Language B, the degree of submissiveness increases when the pitch range is raised; while in Language A, there is no increase of the degree of submissiveness when the pitch range is raised.

c. Type 3 difference: In Language B, the degree of submissiveness increases when the pitch range is raised; while in Language A, the degree of submissiveness decreases when the pitch range is raised.

Because of the universality of the Frequency Code, we think it unlikely that differences of Type 2 or Type 3 will ever be found between any two languages, and we therefore here only consider the question whether there is a Type 1 difference between Dutch and BrE. In other words, is it the case that at identical pitch ranges, one language may express a higher degree of meaning than the other language? The suggestion that there might be a Type 1 difference is fed by the early finding that the two languages in fact do have different standard pitch ranges. According to de Pijper [6], the standard pitch range of Dutch is approximately 70 Hz and that of BrE is approximately 100 Hz. However, it is not clear how the difference in standard pitch range between two languages will reflect differences in degree in the use of the Frequency Code. We can see three possibilities. In situation *a*, listeners from both languages make use of an absolute scale, such that a given F0 value will have an identical value on the Frequency Code, as illustrated in Figure 2a. Second, in situation *b*, listeners project their standard pitch range onto the Frequency Code scale in a relative way, such that a given F0 value will get a higher value on the Frequency Code in Dutch than in BrE, as shown in Figure 2b. Third, in situation *c*, we are assuming that the narrow-range language, Dutch, employs a narrow range exactly because it uses the Frequency Code less intensively. The idea is that since there is not much to be expressed by means of pitch range variation, there is no point in having a wide standard pitch range. By contrast, the wide-range language, BrE, has a wide range exactly because its speakers are much keener to express the meanings encoded in the Frequency Code; hence the wide standard pitch range. Consequently, Dutch listeners are less sensitive to the Frequency Code, so that a given F0 value will have a smaller value on the Frequency Code in Dutch than in BrE. We indicate this as the Use-it-or-lose-it scale, as illustrated in Figure 2c.

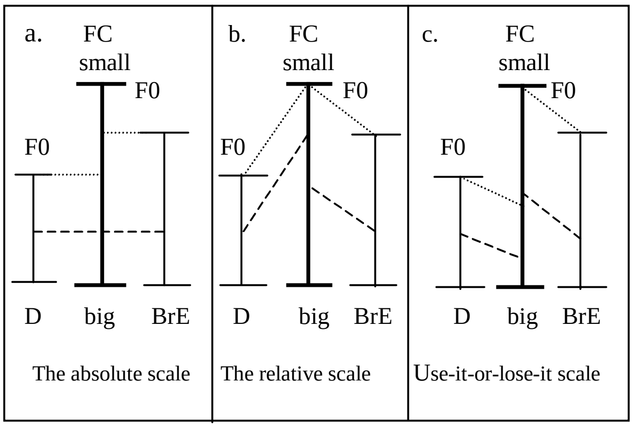

Figure 2. Three implementations of a Type 1 difference. The middle scale is a semantic scale derived from the Frequency code, in which 'small' refers to the meanings associated with high pitch, such as 'dependent', 'appealing', 'uncertain' etc. and 'big' refers to the meanings associated with low pitch, such as 'assertive', 'powerful' and 'authoritative'. The scale on the left of the Frequency Code scale stands for the standard pitch range of Dutch and the scale on the right of the Frequency Code scale stands for the standard pitch range of BrE. The dotted lines indicate how listeners project the standard pitch range onto the Frequency Code scale. The dashed lines indicate the value that a given F0 value corresponds with on the Frequency Code scale.

The question we are interested in thus is which of the three scale types is used by Dutch and BrE. It was suggested in Rietveld et al. [1] that the chosen pitch range might have appeared to be strikingly high to Dutch listeners but normal to BrE listeners. That is, a given pitch range variation seems to be perceived in a relative way. This would mean that the relative scale (see Figure 2b) might be operated in the use of the Frequency Code.

In order to study how pitch range variation affects the use of the Frequency Code, we included two semantic scales that can be derived from the Frequency Code: CONFIDENT vs. NOT CONFIDENT; FRIENDLY vs. NOT FRIENDLY. Our hypothesis is that (1) the degree of confidence will decrease and the degree of friendliness will increase when the pitch range is increased; and that (2) at identical pitch ranges, Dutch will be perceived as less confident but friendlier than BrE. The predicted results are given in Figure 3a and Figure 3b.

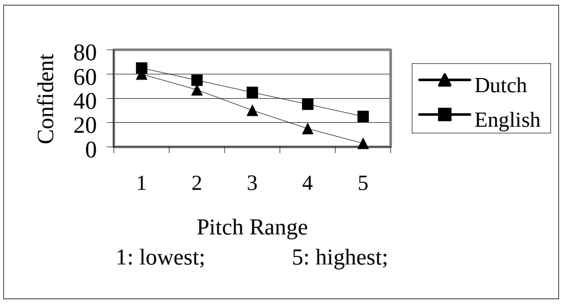

Figure 3a. Predicted results for the scale CONFIDENT vs. NOT CONFIDENT

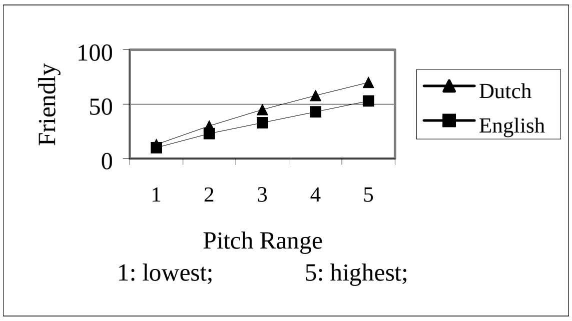

Figure 3b. Predicted results for the scale FRIENDLY vs. NOT FRIENDLY

2. Perception Experiment

In order to show that there is a difference in the use of the Frequency Code between Dutch and BrE, a perception experiment was carried out, in which native speakers from the two languages were asked to listen to a number of sentences in their native language that were varied in pitch range, and to judge these on the semantic scales, CONFIDENT vs. NOT CONFIDENT and FRIENDLY vs. NOT FRIENDLY.

2.1. Stimuli

Since contextual factors such as sentence type, speaker-listener relationship, lexical content of preceding utterance and the intonation of the preceding utterance, may interact with the meaning of intonation contours [7], we designed 12 source utterances exemplifying each of the three speech acts INFORMATION, REQUEST and INSTRUCTION. We used individual sentences to represent each speech act instead of short dialogues. The speech acts were implemented by means of wh-interrogatives, yes-no interrogatives and declaratives, respectively. Examples of such sentences are given below.

INFORMATION: What's the level of this course?
REQUEST: Could you please carry that table down the stairs?
INSTRUCTION: You should fill out the claim form.

Natural productions of the 12 sentences in each language served as the source utterances for the stimuli. They were read by a female Dutch-BrE bilingual speaker. The recording was done in the soundproof studio of the University of Nijmegen and selected utterances were digitized at a 32 kHz rate. Speech manipulation was performed by means of the PSOLA technique.

Each source utterance contained a single accented word, which was assigned both H*LL% and L*HH%. These pitch contours are naturally oriented towards different ends of the

Frequency Code, with the rising L*HH% associating with the 'small' meanings and the falling H*LL% with the 'big' meanings. The stylized versions of H*LL% and L*HH% are given in Figure 4.

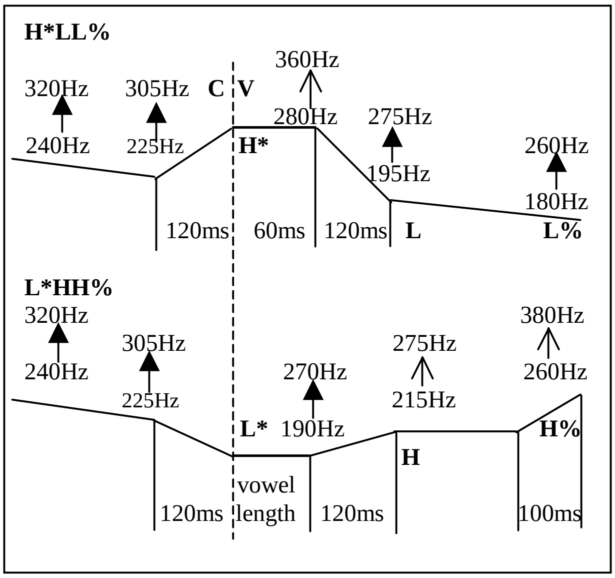

*Figure 4. Stylized contours of H*LL% and L*HH%.*

When considering the question how variation in pitch range is used to signal the meanings encoded in the Frequency Code, one quickly runs into the two ways in which pitch range can be varied: 'register' vs. 'key' [7,8] or 'overall level' vs. 'span' [7,8]. 'Register' and 'overall level' refer to the average F0 of a stretch of speech, indicated as pitch register here. 'Key' and 'span' refer to the difference between the highest F0 and the lowest F0 across a stretch of speech, indicated as pitch span here. We varied the contours both in pitch register and pitch span, as they may affect the Frequency Code in different ways. With regard to variations in pitch register, both H-tones (H*, H, H%) and L-tones (L*, L, L%) were varied in equal steps, as indicated by both the solid and the plain upper arrows in Figure 4. This gave us 120 stimuli (3 speech acts × 4 sentences × 2 pitch contours × 5 steps) varying in pitch register in each language, referred to as the Pitch Register stimuli. With regard to variations in pitch span, only the H-tones in each contour were varied, while the L-tones remained the same, as indicated by the upper arrows in Figure 4. This gave us 120 (3 speech acts × 4 sentences × 2 pitch contours × 5 steps) stimuli varying in pitch span in each language, referred to as the Pitch Span stimuli. The maximum difference between the highest F0 and the lowest F0 is 140 Hz in H*LL% and 190 Hz in L*HH%. In this way, the standard pitch range of BrE is covered as well as that of Dutch [6].

In order to minimize the risk that subjects would detect a pattern of pitch range variation in specific pitch contours, we added 128 fillers to the experimental stimuli in each language. The experimental stimuli and the fillers were mixed randomly and then divided into 23 blocks of 16 stimuli in each language. To provide listeners with a point of orientation on the scale, we added an anchor stimulus to the beginning of each block of stimuli. There was a 4.5 s pause between stimuli and a 7 s pause between blocks. Each block was preceded by a 300 Hz sine wave warning tone. In order to avoid an order effect, we prepared four stimulus orders by randomizing each block internally and the blocks as a whole four times. Each of the fours stimulus orders was recorded onto a TDK audiotape. This gave us four 43-minute stimulus tapes in each language.

In addition, we chose to use the magnitude estimation method [9] to obtain the perceptual judgments as in Rietveld et al. [1].

2.2. Procedure

Fifty-three linguistically naïve native speakers of Dutch and 51 of BrE between 18 and 30 years old took part in the experiment in equivalent circumstances. Subjects were instructed to try to imagine themselves as the addressees and indicate what kind of impression the speaker made on them by drawing a slash on a 100mm scale: the more to the left the slash was placed, the more 'NOT CONFIDENT' or 'NOT FRIENDLY', and the more to the right the slash was placed, the more 'CONFIDENT' or 'FRIENDLY' the sentence sounded.

3. Results

Data obtained from the experiment were divided into two sets, (1) the Pitch Register data and (2) the Pitch Span data. Two Analyses of Variance were performed on the Pitch Register data and another two Analyses of Variance were performed on the Pitch Span data, for each of the two dependent variables: the scale CONFIDENT vs. NOT CONFIDENT and the scale FRIENDLY vs. NOT FRIENDLY. The analyses comprised one between-subject factor: Language, and three within-subject factors: Pitch Contour (2 levels), Speech Act (3 levels), and Pitch Register (5 levels) for the Pitch Register data or Pitch Span (5 levels) for the Pitch Span data. We adopt a significance level of 0.05 and report p-values after correction for sphericity (Huynh-Feldt's ε) [10]. As we are mainly interested in the effects of pitch range variation on the interpretation of the pitch contours H*LL% and L*HH% as a function of Language in Dutch and BrE, we will here only consider significant interactions concerning the between-subject factor Language and the within-subject factors Pitch Contour, Pitch Register and Pitch Span.

3.1 The scale CONFIDENT vs. NOT CONFIDENT

With regard to the Pitch Register data, we found a significant interaction of Pitch Contour × Pitch Register ($F_{4,248}$=15.61, $p<0.001$) and a significant interaction of Language × Speech Act × Pitch Register ($F_{8,496}$=3.70, $p<0.001$). For the Pitch Span data, there was a significant interaction of Pitch Contour × Pitch Span ($F_{4.248}$=3.68, $p<0.001$). The two-way interaction of Pitch Contour × Pitch Register and that of Pitch Contour × Pitch Span is such that the degree of confidence decreases when the pitch register or the pitch span rises. This is apparently a manifestation of the Frequency code: higher pitch sounds vulnerable and submissive; hence a contour may sound less confident at a higher pitch. Furthermore, by and large, H*LL% is perceived as more confident than L*HH%. This is also a manifestation of the Frequency Code: low pitch sounds protective and dominant; hence a low-ending contour may sound more assertive and more confident. Since our main concern is not the effect of Speech Act, we pooled the results of the interaction of Language × Speech Act × Pitch Register over the speech acts. By this means, we found that although the general tendency is that the higher the pitch register is, the lower the degree of confidence is, at an identical pitch register, BrE is perceived as more confident than Dutch, in particular at higher levels of pitch register, as shown in Figure 5.

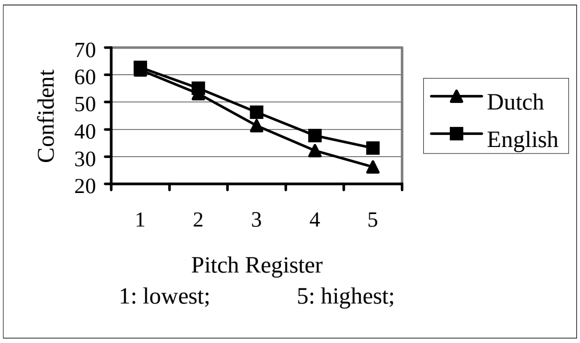

Figure 5. The interaction of Language × Speech Act × Pitch Register pooled over speech acts.

3.2 The scale FRIENDLY VS. NOT FRIENDLY

With respect to the Pitch Register data, we found a significant interaction of Pitch Contour × Pitch Register ($F_{4,248}=3.67$, $p<0.007$) and a significant interaction of Language × Pitch Register ($F_{4,248}=3.90$, $p<0.036$). For the Pitch Span data, there was a significant interaction of Pitch Contour × Pitch Span ($F_{4,248}=2.84$, $p<0.025$). The two-way interaction of Pitch Contour × Pitch Register and that of Pitch Contour × Pitch Span is such that the degree of friendliness increases when the pitch register or the pitch span is raised. Moreover, by and large, L*HH% is perceived as friendlier than H*LL%. These results are again manifestations of the Frequency Code. The two-way interaction of Language × Pitch Register is such that at an identical pitch register, BrE is perceived as friendlier than Dutch by the respective native speakers, as shown in Figure 6.

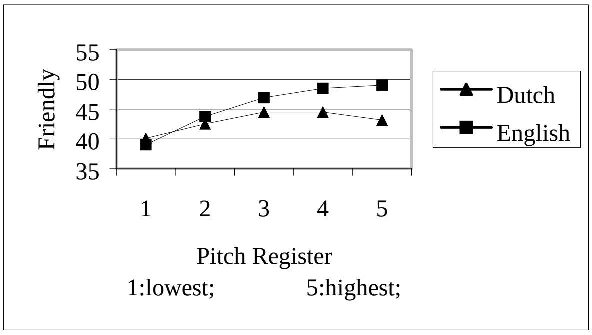

Figure 6. The interaction of Language × Pitch Register.

4. Conclusions

Although by and large, H*LL% is the more-confident-sounding contour and L*HH% is the friendlier-sounding contour, in both contours, the degree of confidence decreases and the degree of friendliness increases when the pitch range is increased. This is true both for Dutch and BrE. This reflects the effects of pitch range on the perception of the universal meaning, embodied in the Frequency Code. Moreover, at identical pitch ranges, BrE is perceived as more confident and friendlier than Dutch. This is a difference in degree of the use of the Frequency Code, and hence represents a language-specific meaning component. That BrE is perceived as more confident than Dutch is in accordance with our hypothesis (see Figure 3a and Figure 5) based on the assumption of the relative scale (see Figure 2b): a given pitch range is perceived as higher in Dutch than in BrE due to the smaller standard pitch range of Dutch and therefore is perceived as less confident in Dutch than in BrE. However, the fact that BrE is perceived as friendlier than Dutch cannot be accounted for by the relative scale, but can only be explained by the Use-it-or-lose-it scale. According to the relative scale, a given pitch range would be perceived as higher in Dutch and therefore as friendlier in Dutch than in BrE. According to the Use-it-or-lose-it scale, Dutch speakers do not use the Frequency Code as intensively and consequently Dutch speakers are less sensitive to the Frequency Code. Therefore, a given pitch range will score lower on semantic scales derived from the Frequency Code. In the case of FRIENDLY vs. NOT FRIENDLY, a given pitch range will be perceived as less friendly in Dutch than in BrE. As the Use-it-or-lose-it scale can also be used to account for the fact that BrE is perceived as more confident than Dutch, it is likely that the Use-it-or-lose-it scale is used in Dutch and BrE. This scenario, therefore, suggests that the reason why Dutch has a smaller pitch range than English is that its speakers are less concerned to signal universal meanings expressed by pitch range variation, like the Frequency Code, than speakers of BrE.

5. References

[1] Rietveld, A., Gussenhoven, C., Wichmann, A., and Grabe, E., "Communicative effects of rising and falling pitch accents in British BrE and Dutch," *Proceedings of the ESCA workshop on dialogue and prosody.*, 111-116, 1999.

[2] Ohala, J.J., "Cross-language use of pitch: an ethological view", *Phonetica* 40: 1-18, 1983.

[3] Ohala, J.J., "An ethological perspective on common cross-language utilization of F0 in voice," *Phonetica* 41: 1-16, 1984.

[4] Ohala, J.J., "The frequency code underlines the sound symbolic use of voice of pitch". In Hinton, L. et al. (eds.) *Sound symbolism.* Cambridge: Cambridge University Press, 1994.

[5] Gussenhoven, C. and Chen, A.J., "Universal and language-specific effects in the perception of question intonation," *Proceedings of the 6th ICSLP.*, II: 91-94, 2000.

[6] de Pijper, J., *Modelling British English Intonation*, Dordrecht: Foris Publications, 1983.

[7] Cruttenden, A., *Intonation.* Cambridge: Cambridge University Press, 1986.

[8] Ladd. D.R., *Intonational Phonology.* Cambridge: Cambridge University Press, 1996.

[9] Zraick R. and Liss, J.M., "A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality," *Journal of Speech, Language, and Hearing Research* 43: 979-988, 2000.

[10] Rietveld, A. and van Hout, R., *Statistical techniques for the study of language and language behaviour*. Berlin: Monton de Gruyter, 1993.

Acknowledgements: We would like to thank Gerard Latjes, Eric Kellerman, Brechtje Post and Patricia Vermillion for their help with the administration of the tests, Hedy Kamara, Rob Kerkhoff, Ellen Kerkhoff and Sofie Kienhorst for assistance with the recording and processing the data, and Joop Kerkhoff for allowing us to freely call upon his programming expertise.