

Belfast Intonation in L2 speech

Christiane Ulbrich

School of Communication & Social and Policy Research Institute, University of Ulster, Belfast

C.Ulbrich@ulster.ac.uk

Abstract

The present paper investigates the production and perception of rising intonation patterns in first language (L1) speakers of Belfast English and native German learners of English with and without previous exposure to the Belfast variety of English (BfE). Whilst there is evidence that Northern Standard German (NSG) predominantly uses falling nuclear pitch patterns in declaratives, Swiss German (SG) esp. the variety spoken in Bern [6] and BfE were previously found to produce mainly rising pitch patterns in nuclear position of declaratives. The paper investigates the question if rising pitch patterns produced by SG speakers are transferred into their L2 BfE and if so, do these cross-language similarities result in different ratings of foreign accent compared to NSG speakers. Thus two issues are addressed: (i) target association vs. target alignment and (ii) the effect of cross-varietal differences in L1 on the success of L2 acquisition.

1. Introduction

Current research in L2 acquisition rests on two rather tentative assumptions: (i) although there is a common agreement that L1 plays a role in L2 acquisition [5, 9, 23, 24] and that its characteristics can be facilitating as well as hindering in the acquisition process of an L2, no formal account has so far been provided to explain the influence of the native grammar on the acquisition of an L2. Secondly, whilst the role of prosodic characteristics in L2 acquisition has been widely acknowledged in the past their distinct nature from phonetic and phonological properties of segmental characteristics has not been integrated into current models of L2 acquisition (e.g. [7] Speech Learning Model (SLM) [8] or the Perceptual Assimilation Model (PAM) [1, 2]). These still need to be revised in order to equally account for the segmental and the prosodic level of speech production and perception as well as their interaction. Some recent studies have focused on the contribution of prosodic characteristics such as intonation, speaking rate, timing and rhythmic patterns [e.g. 14, 16, 17, 20, 22, 4] to the perception of a foreign accent (FA).

A previous study, also based on recordings of native German speakers acquiring L2 BfE, has shown that regional characteristics – on both, the segmental and the prosodic level – are acquired by L2 learners [21] and that those characteristics contribute to a lower rating of foreign accentedness [22]. The German group was divided into speakers with long term exposure to BfE and speakers without previous exposure to BfE. A comparison of the two groups showed that German speakers with long term exposure had acquired rising intonation patterns, which are regionally marked in BfE [11, 21], in nuclear accent position of declaratives. On the segmental level we found some regionally marked vowel realizations and post-vocalic /r/ [21]. The results strongly suggested that segments appear to play a greater role in foreign accent rating (FAR) than prosody does (as was also found in [4]). Nevertheless, intonation also

appeared to contribute to the perception of FA. By using a procedure of prosodic transplantation it was found that crossing native BfE segments with L2 intonation resulted in a higher FAR. An even higher FAR was scored when BfE segments were crossed with intonation patterns of German speakers that had no previous exposure to BfE. The results led to the conclusion that L2 learners can acquire a fairly native-like intonation. This was confirmed by the fact that stimuli with both segments and intonation of German speakers (with no previous exposure to BfE) got equally lower FAR when the intonation was swapped with intonation patterns of either BfE or L2 speakers of English with extensive exposure to the regional variety spoken in Belfast.

The extent of transfer from L1 to L2 varies depending on many variables. Relevant are at least the following: stage and mode of learning, input, social setting and individual learner features (e.g. aptitude, age, metalinguistics awareness, literacy, language proficiency, education, social background etc.). The complex interaction of transfer with these variables is still not well understood. Evidence of L1 transfer is clearly seen in the area of phonology. It has, however, mainly been shown to be negative transfer, its positive effects have not been studied to the same extent. An ongoing debate in the current literature relates to the question of transfer depending on the proximity between L1 and L2. Some of the results suggest that the acquisition of L2 sounds and/or structures that are non-existent or vary considerably from those of L1 can be less erroneous and therefore easier to acquire than those that are formally rather similar [18].

This issue will be addressed in the present paper by comparing stimuli containing transplanted SG prosody with copy-resynthesized stimuli produced by German native speakers that have not been exposed to BfE. The results will give insight into the relative importance of phonological realization of nuclear pitch accents versus their phonetic implementation thus investigate target association and alignment in the synchronization between the segmental and the prosodic level of speech. Additionally, we will address issues of cross-language and cross-varietal prosodic typology. Lastly, a comparison of FAR for those two groups of speakers might shed light onto the question if similarities in the form of pitch pattern in L1 and L2 allow L2 learners to acquire more easily a native-like intonation.

2. Method

The present paper is based on a corpus of read speech produced by four groups of speakers:

- native speakers of Belfast English (B)
- native Swiss German speakers of the Bernese dialect with no previous exposure to BfE (SG)
- native German speakers of NSG who have lived in Belfast for a minimum of three years (L2)
- native German speakers of NSG with no previous exposure to BfE (G).

A total number of 20 female speakers (five per group) was recorded in a quiet room using a Sennheiser ME64 directional condenser microphone (cardioid, frequency response 40-20,000 Hz, ± 2.5 dB) with a sampling rate of 22,050 Hz as WAV files directly onto a Toshiba notebook computer for processing and analysis in PRAAT [3]. The statistical analysis was done using SPSS. The reading task was carried out by all groups of speakers in English and additionally in German by the native speakers of German. Ten declarative utterances, each containing one of ten targets (five disyllabic and five trisyllabic words), were embedded into a longer text always at the end of a short paragraph. This was done to avoid intonation contours indicating continuation, which have been shown to differ considerably (i) between regional varieties (e.g. [10]) and (ii) from pitch patterns indicating termination (e.g. [12]). The targets are cross-linguistically comparable regarding segmental content and stress placement and they appeared in nuclear position of short, broad focus utterances of comparable length. Two examples are given below.

Example 1: Context: Preparation of a fruitsalat
 German: *Er nahm die Mángo.*
 English: *He took the Mángo.*

Example 2: Context: Talking about last night.
 German: *Sie sahen ein Vídeo.*
 English: *They watched a vídeo.*

The total of 350 target sentences was extracted for further analysis. The corpus was segmented and phonemically transcribed by a phonetician using the IPA based on perception and visual clues provided in spectrograms. The nuclear pitch accents were intonationally labeled using the rhythmic tier and the phonetic tier in an adaptation of the IViE system [11]. A PRAAT script was employed to measure f_0 (st) and duration (ms) at two (potentially three) points within each voiced portion of each phoneme transcribed: at the beginning, the end and at potentially appearing f_0 turning points in the pitch contour. The measurements were manually inspected and subsequently returned in separate tiers of the textgrid for the analysis of f_0 movement and alignment and in preparation of the stimuli creation for the perception task. The following two subsections detail the methods employed in the production task and the perception experiment.

2.1. Production task

The production task was carried out in order to compare the realization of nuclear pitch patterns (i) across the four subject groups and (ii) for the native speakers of German across the two languages. We compared the realization of pitch accent patterns (high target, low target, rising and falling contour) in German for the three groups of native German speakers and in English for all groups of speakers (including the B speakers). Resulting from the phonetic labeling of the pitch patterns we distinguished the following realizations:

High pitch accents	(H)
Low pitch accents	(L)
Rising pitch accents	(LH)
Falling pitch accents	(HL)
Rising-Falling pitch accents	(LHL)

Given the purpose of the study we were not interested in a phonological typology of high or low targeting pitch accents and boundary tone realizations in the sense of the autosegmental-metrical approach [19]. Therefore the material was not designed in a way to distinguish between for example falling patterns (H*L) followed by a low boundary tone (L%) or a non-specified boundary tone (%). The analysis of the production task was effectively carried out to confirm cross-

varietal and cross-language differences in our corpus that have previously been found and described in the literature [e.g. 6] in order to establish our hypothesis. We were more interested in a general directionality of pitch patterns, the actual realization of high or low pitch within the accented syllable, the pitch contour following (as indicated by the labels above) and the alignment of the f_0 -contour within the accented syllable. The prosodic annotation of nuclear pitch patterns was carried out by four trained annotators, blind to conditions. In order to test for reliability of annotations across the four annotators we compared the annotations. 78% of the scores were consistent across annotators. The majority of errors 14% were found in annotations of rising patterns (LH vs. LHL). An extended version considering segmental and phonotactic content of target words and their interaction with prosodic boundaries would allow for a more detailed analysis of synchronization effect between segmental level and intonation contour.

2.2. Perception experiment

The perception experiment was based on English utterances only. The creation of stimuli for the perception experiment employed a procedure known as prosodic transplantation. This method uses the PSOLA algorithm and has already successfully been employed in the study of FA perception [e.g. 4, 22]. By means of re-synthesis extracted pitch and duration values of one voice are transplanted onto segmentally identical units of another voice. The target sentences of one native B speaker and one G were selected. Their original pitch and duration values were then replaced by those produced by SG and L2 speakers at two (potentially three) points within each syllable. In order to produce stimuli with a comparable sound quality all utterances including those produced by the five B speakers and the five G speakers were copy-synthesized.

The stimuli were presented via speakers to 48 native speakers of BfE; students at the University of Ulster all with normal hearing and unpaid for their participation in the experiment. The participants were not urged to answer immediately but they could listen to each stimulus only three times. During the instructions participants were informed that they would listen to acoustically modified stimuli. They were asked to judge each stimulus by indicating if the sentence they just heard was produced by a native or non-native speaker of BfE. Following this forced-choice paradigm we asked for a confidence rating of their choice on a 3-point scale (certain, semi-certain, and uncertain) which resulted in an operational 6-point scale of FAR. Based on a previous study [22] we established the following hypothesis (referring to the conditions provided in table 1):

- Condition 1: highest FAR
- Condition 8: lowest FAR
- Condition 2: FAR between Cond.1 and Cond.8

In addition we predicted that a comparison between Condition 1 vs. Condition 3 would give insight into the relative importance of phonological or phonetic realization of pitch accents in FAR. If FAR for Cond.3 < Cond.1 then the realization of the same phonological pitch pattern, thus the target association could be sufficient to lower the perception of FA. If FAR for Cond.3 = Cond.1 then the phonetic implementation, thus the target alignment could be the more important cue for FA perception. A comparison of segmental strings produced by G and BfE speakers with prosodic characteristics of L2 and SG speakers (cond. 4&5 and cond. 7&8) would provide more insight into the relative contribution of phonetic implementation of pitch accent patterns to the FA

perception and provide evidence that regionally marked subtle phonetic details can only be acquired in a naturalistic setting by L2 learners.

3. Results

In the first part of this section we will present the results of the production task and their contribution to the established hypothesis and predictions for the perception experiment.

3.1. Results of the production task

The results of the production task showed (i) cross-varietal differences in the realization of pitch patterns between SG and NSG speakers in the German stimuli and (ii) cross-language differences between all four groups of speakers (including B) in the English stimuli.

Table 1: Number of stimuli for the perception test (*S*=Segments; *P*=Prosody, *CR*=Copy-resynthesis)

Condition	Prosodic Transplant	Produced utterances	Perception
1 CR	G- <i>S</i> G- <i>P</i>	50	43
2 CR	L2- <i>S</i> L2- <i>P</i>	50	27
3 CR	SG- <i>S</i> SG- <i>P</i>	50	27
4	G- <i>S</i> L2- <i>P</i>	50	43
5	G- <i>S</i> SG- <i>P</i>	50	43
6 CR	B- <i>S</i> B- <i>P</i>	50	43
7	B- <i>S</i> L2- <i>P</i>	50	43
8	B- <i>S</i> SG- <i>P</i>	50	43

3.1.1. German stimuli

Differences in the realization of nuclear pitch patterns between the three groups of L1 German speakers showed that SG speakers realized more rising pitch patterns compared to both groups of NSG speakers. We found that SG speaker predominantly produced low, rising and rising-falling pitch accents whereas G speakers produced mainly falling pitch accents. Interestingly, L2 speakers produced also more rising-falling pitch accents compared to the G group but still produced mainly falling pitch accents in nuclear accents. The results are presented in table 2.

Table 2: Mean frequencies of pitch patterns in 25 bi-syllabic (left) and 25 tri-syllabic (right) German target words

	H	HL	L	LH	LHL
G	2	19 22	4 1		2
L2		15 18	2	3	5 7
SG	3	1	2	10 4	9 21

3.1.2. English stimuli

The analysis of the English production task shows similar results for G (mainly HL). However, the L2 speakers produced considerably more rising (LH) and rising-falling (LHL) pitch accents compared to their German productions indicating that

they have acquired regionally marked pitch patterns. Interesting here is the realization of considerably more falling (HL) pitch patterns by the SG speakers compared to their realizations in the German production task. The L1 speakers of Belfast English realized predominantly rising and rising-falling pitch accents, however, one third of the utterances was produced with falling pitch accents which confirms previous findings [21] that speakers tend to shift their pronunciation in more formal speech towards a perceived standard in the present case the Southern British variety of English (SBE) a variety that features predominantly falling pitch accents in nuclear accents of declaratives [11] (see table 3).

Table 3: Mean frequencies of pitch patterns in 25 bi-syllabic (left) and 25 tri-syllabic (right) English target words

	H	HL	L	LH	LHL
G	2	23 24			1
L2		11 10	2	9 3	3 12
SG		12 11		8	5 14
B		7 7		8 9	10 9

3.2. Results of the perception experiment

Both, SG and L2 speakers produced 27 targets with rising or rising-falling pitch patterns; typically found in BfE. Only those utterances underwent prosodic transplantation with the segmental material of the same utterance produced by one G and one B speakers. The same target utterances were copy-resynthesized for all four groups of speakers. Since SG and L2 speakers did not realize the exact same utterances with a rising or rising-falling pattern (8 were different) the total number of test items was 232. These were presented to 48 native speakers of BfE according to the method detailed in section 2.2. The experiment took place in a quiet room and lasted about 40 minutes. The stimuli were presented in quasi-random order and the actual test was preceded by a trial of five stimuli to ensure familiarity with the task. The obtained scores for FAR were submitted to an overall ANOVA (GLM, repeated measurements) in SPSS with FAR score as dependent variable. The overall group effect was highly significant ($F_{(7,1128)}=7963.801$). A post-hoc test (Turkey HSD) revealed significant differences between all 8 conditions with the exception of cond. 1 and 3. The difference between copy-resynthesized L2 and SG on the one hand and their transplanted versions on the other was also significant. Additionally, we found that transplantation of both L2 and SG prosody onto G segments resulted in a significant decrease of FAR for G segments and in a significantly increased FAR for B segments. A comparison of FAR for condition 1, 6 and 2 confirmed our hypothesis in that copy-resynthesized G (cond. 1; FAR 5.3) received the highest FAR, B (cond.6; FAR 1.05) received the lowest FAR and L2 (cond. 2; FAR 2.92) received a FAR in between the two. Note that the FAR score for G and SG do not differ significantly. Transplanting prosodic characteristics of L2 onto G (cond. 4) yielded a lower FAR (4.74) compared to the copy-resynthesis of G. Also, a lower FAR (3.55) was scored for stimuli of G segments with SG prosody (cond. 5). The comparison of FAR for copy-resynthesized B with those of condition 7 (B segments and L2 prosody) and condition 8 (B segments and SG prosody) shows that although FAR scores are lower compared to copy-resynthesis of L2 and SG, they are still higher compared to scores obtained for 'original' B utterances (cond. 7 FAR 1.24; cond. 8 a FAR of 2.39).

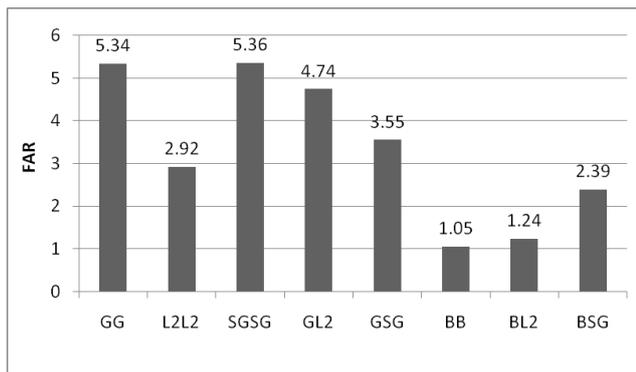


Figure 1: FAR score obtained in the perception test

4. Discussion and Conclusions

Our results suggest that cross-language similarities in the realization of nuclear pitch patterns are (partly) transferred into L2. The realization of more falling pitch patterns by SG speakers in the English production task accent might point to an effect of class room instructions. The variety of English taught in (Swiss) German education is based on SBE known to feature falling accents in nuclear position of declarative utterances. The comparison of pitch patterns realized by L2 speakers in the German and English stimuli could be interpreted as transfer from L2 to L1. It appears that compared to G speakers L2 speakers produced more rising and rising-falling patterns in their German L1, a pitch accent realization seen relatively seldom in NSG in nuclear position of declaratives.

The results of the perception experiment replicated results previously found in [22]. Additionally, they show that prosodic similarities between L1 and L2 facilitate the acquisition of L2 prosody. In the present case similar realization of rising and rising-falling pitch patterns in SG and BfE suggest that similarities in the target association between L1 and L2 contribute to a lesser degree of FA. However, a comparison of scores obtained for L2 and SG speakers suggest that there are more subtle prosodic cues to be acquired than mere target association and gross pitch movement in order to acquire a native-like L2 prosody. These prosodic cues might be found either in the alignment of pitch targets to the segmental string in which case FA might rather be seen as a result of erroneous phonetic implementation. The differences could however also be caused by prosodic characteristics preceding the actual nuclear pitch accent, a question that needs to be addressed in more detail. The similar scores for G and SG copy-resynthesized stimuli suggests that segmental characteristics act as ‘give-away’ for FA since the extractions of SG prosodic characteristics yielded a lower FA when transplanted onto G. In summary, regarding typological similarities [e.g. 15, 13] it seems that ‘systemic’ differences, relating to differences in the inventory of phonologically distinct pitch patterns facilitate L2 acquisition and thereby result in a less ‘strong’ FA. However, there are additional ‘realizational’ and ‘phonotactic’ differences (details in the phonetic realization of what may be phonologically the same pattern and its permitted structure) that need to be further investigated. Therefore, further research needs to address more global aspects pertaining to syllables, prosodic words and intonational phrases and also study potential interaction effects between prosodic characteristics and the segmental string by employing a larger and more carefully controlled corpus.

5. References

- [1] Best, C., “A direct realist view of crosslanguage speech perception”, in W. Strange [Ed.], *Speech perception and linguistic experience: Issues in cross-language research*, Timonium, MD: York Press, 171-204, 1995.
- [2] Best, C. T., McRoberts, G. W. and Goodell, E., “Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listeners’ native phonological system”, *JASA* 109(2): 775-792, 2001.
- [3] Boersma, P. and D. Weenink, E., *Praat: doing phonetics by computer (Version 5120)*. 2005.
- [4] Boula De Mareuil, P. and Vieru-Dimulescu, B., “The contribution of prosody to the perception of foreign accent”, *Phonetica*, 63(4): 247-267, 2006.
- [5] Brière, E., “An investigation of phonological interferences”, *Language* 42:768-95, 1966.
- [6] Fitzpatrick-Cole, J., “The alpine intonation of Bern Swiss German” *ICPHS*. 941-944, 1999.
- [7] Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H. and Tsukada, K., “Degree of foreign accent in English sentences produced by Korean children and adults”, *Journal of Phonetics*, 34:153-175, 2006.
- [8] Flege, J., “Second language speech learning: Theory, findings, and problems”, in W. Strange [Ed.], *Speech Perception and Linguistic Experience: Issues in Crosslanguage Research*, 233-73. Baltimore: York Press. 1995.
- [9] Flege, J., “The phonological basis of foreign accent”, *TESOL Quarterly* 15: 443-55, 1981.
- [10] Gilles, P., „Regionale Prosodie im Deutschen Variabilität in der Intonation von Abschluss und Weiterweisung”, *Linguistik - Impulse & Tendenzen* 6. 2005.
- [11] Grabe, E., “Intonational variation in urban dialects of English spoken in the British Isles”, in P. Gilles and J. Peters [Eds.], *Regional Variation in Intonation*. Linguistische Arbeiten, Tuebingen, Niemeyer, 9-31, 2004.
- [12] Grice, M. and Bauman, S., “An Introduction to Intonation – Functions and Models”, in J. Trouvain and G. Ulrike [Eds], *Non-native Prosody: Phonetic Description and Teaching Practice*, 25-52, Mouton de Gruyter, 2007.
- [13] Hirst D. and Di Cristo, A., “Intonation Systems. A Survey of Twenty Languages”, Cambridge: Cambridge University Press, 1998
- [14] Jilka, M. “The contribution of intonation to the perception of foreign accent”, PhD thesis, University of Stuttgart. 2000.
- [15] Jun, S-A. “Prosodic Typology, The Phonology of Intonation and Phrasing”, Oxford, Oxford University Press. 2006.
- [16] Magen, H., “The perception of foreign-accented speech”, *Journal of Phonetics*, 26:381-400, 1998.
- [17] Munro, M.J., “Nonsegmental factors in foreign accent: Ratings of filtered speech”, *SSLA*, 17:17-34, 1995.
- [18] Odlin, T., *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge University Press, 1989.
- [19] Pierrehumbert, J. “*The Phonology and Phonetics of English Intonation*.” Ph.D. dissertation, MIT. 1980.
- [20] Trofimovich, P. and Baker, W., “Learning second language suprasegmentals: Effect if L2 experience on prosody and fluency characteristics of L2 speech”, *Studies in Second Language Acquisition*, 28:1-30, 2006.
- [21] Ulbrich, C., “Acquisition of regional pitch patterns in L2”, In *Speech Prosody 2008*, 575-578, 2008.
- [22] Ulbrich, C. and Mennen, I. (in prep.): *Dissociation of segments and intonation in L2. Bilingualism: Language and Cognition*. Cambridge University Press. Cambridge.
- [23] Wode, H., “Categorical perception and segmental coding in the ontogeny of sound systems: A universal approach” in C. Ferguson, L. Menn and C. Stoel-Gammon [Eds.], *Phonological Development: Models, Research, Implications*, 605-31, Baltimore: York Press, 1992.
- [24] Wode, H., “The beginnings of L2-phonological acquisition”, *IRAL* 16:109-24, 1978.