

# German pitches in English: Production and perception of cross-varietal differences in L2\*

CHRISTIANE ULBRICH  
Philipps University Marburg, Germany  
& University of Ulster, Northern Ireland

(Received: October 18, 2010; final revision received: October 17, 2011; accepted: September 19, 2012; first published online 20 December 2012)

*The present study examines the effect of cross-varietal prosodic characteristics of two German varieties, Northern Standard German (NG) and Swiss German (SG), on the production and perception of foreign accent in L2 Belfast English. The analysis of production data revealed differences in the realisation of nuclear pitch accents in L1 German and L2 English produced by the two groups of speakers. Foreign accent ratings of L2 English sentences produced by NG speakers with no or extensive experience, native Belfast English speakers and SG speakers were obtained from native Belfast English listeners. The findings showed that segmental and prosodic characteristics play a role in the perception of foreign accent. In addition, they can be more similar across languages than across varieties of the same language. This in turn affects which and how cross-varietal differences in L1 impact on the degree of perceived foreign accentedness. The results are consistent with a usage-based account.*

Keywords: acquisition, prosody, cross-variation, German, English

## 1. Introduction

Over the past few decades, the field of Second Language Acquisition (SLA) has greatly benefited from developments in methodological paradigms and experimental techniques, a growing pool of data, theoretical advances and the application of integrated cross-disciplinary approaches. These of course are the necessary prerequisites allowing scholars working in the field of SLA to gain a deeper understanding of the architecture of the grammar (Jackendoff, 1997) and its bearing on what has to be acquired and what comes for free in first and second language acquisition. With regard to the relationship between phonetics and phonology, modularity has been challenged in numerous publications (Ohala, 1990). Both substantial and subtle variations between groups and individuals in second language acquisition (as well as in other domains, for instance L1 acquisition, pathology, sociophonetics) have led to proposals of more non-deterministic, variable accounts for sound structure (Hazan, 2007; Pierrehumbert, 2001; Scobbie, 2007), since variability is often not accidental or unsystematic but gradient.

The multi-faceted role of prosody adds to the complexity of the relationship between different modules

of the grammar. Prosody has been shown to play an important role in the conveyance of linguistic information (Nespor & Vogel, 1986; Selkirk, 1995; Zubizarreta & Vergnaud, 2005), the regulation and structure of discourse (Pierrehumbert & Hirschberg, 1990; Wichmann, 2000a) as well as the indication of speaker related physical, psychological and social factors (Benus, Gravano & Hirschberg, 2007; Mozziconacci & Hermes, 1998; Wichmann 2000b). Analysis of prosody includes:

- (i) temporal characteristics of duration, speech and articulation rate,
- (ii) tonal characteristics, identified as movements and excursion in fundamental frequency ( $f_0$ ),
- (iii) loudness, measured by overall intensity and spectral intensity,

as well as the contribution of all three to rhythm, intonation and stress.

Most of this research, however, has dealt with prosodic features in isolation, largely detached from the segmental level of speech, the phonemic string. The interaction between segments and prosody has only been considered in typology and cross-linguistic comparisons (Pierrehumbert & Beckmann, 1988). The investigation of alignment and association, i.e. the timing of  $f_0$  peaks and valleys with respect to segmental landmarks, has dominated the field's interest and has led to descriptions of intonational systems (Gussenhoven, 2004; Hirst & Di Cristo, 1998; Jun, 2006; Ladd, 2008). These systems are generally associated with an inventory of prosodic cues or parameters similar to those exploited by the

\* I would like to thank the anonymous reviewers for their helpful and constructive comments that greatly contributed to improving the final version of the paper. This research was supported by the U.K. British Academy (BA) through grant no. SG4501 to University of Ulster. This support is gratefully acknowledged. I thank the University of Ulster for the use of the Phonetics Laboratory, and all subjects for their participation.

Address for correspondence:

Institute of German Linguistics, Philipps University Marburg, Deutschhausstrasse 3, D-35032 Marburg, Germany  
christiane.ulbrich@staff.uni-marburg.de

segmental phonemic systems (for instance vowel systems can be characterized, for example, on the basis of backness and rounding; see Crothers, 1978). Naturally, then, there are systematic cross-linguistic similarities and differences between the prosodic systems of languages (Pierrehumbert, 2000), much like those found on the segmental level (for example voice-onset-time (VOT) in plosives, see Cho & Ladefoged, 1999; Ladefoged & Cho, 2001, or degree of retroflexion, see Ladefoged & Maddieson, 1996).

Additionally, phonetic variability and its impact on language typology has been studied at both the segmental level and the level of prosody. It now becomes crucial to progress from these two relatively well explored yet largely independent areas to the consideration of their complex interaction in speech production and perception. Extensive work in the area of laboratory phonology has focused on the complexity of the incorporation of phonetic gradient variability into systematic phonological structure. More recently, this focus was naturally extended to take into account the intricacy of the interaction between segmental and prosodic levels of speech.

Several phonological and prosodic phenomena such as syllable, word and phrase boundaries or syllable structure have been shown to be directly reflected in the duration and the overlap of articulatory gestures (Bombien, Mooshammer, Hoole, Kühnert & Schneeberg, 2006; Byrd, 1996; Byrd & Choi, 2006; Cho & Keating, 2009; Kim & Cho, 2009). Furthermore, the realisation of prosodic events such as  $f_0$  movement, height or range has been shown to be constrained by articulatory mechanisms that regulate the alignment of separately controlled laryngeal and supra-laryngeal movements (Fujimura, 2000; Krakow, 1999; MacNeilage, 1998), permitting various settings of prosodic and segmental components of speech that in turn contribute to communicative functionality.

The necessity for an integrated approach to the analysis of segments and prosody at the interface between phonetics and phonology is self-explanatory. Understanding of variability and gradience in the segmental and prosodic domains, and their implementation on the phonological level cannot explain surface variation that involves variation resulting from the interaction between the two domains. Such considerations are especially imperative in the area of SLA, since cross-linguistically similar prosodic features aligned with systematically different segmental characteristics or vice versa may produce comparable speech events.

### 1.1 *Foreign accent*

As pointed out by Gut (2007), foreign accent (FA) still lacks a comprehensive and universally accepted definition. Foreign accent is generally described as non-

native speech that deviates from native speech and it is understood as the result of cross-linguistic differences between two (or more) phonological systems. Foreign accent can be analysed from the perspective of production, measuring acoustic phonetic aspects that deviate from native speech in production. However, it can also be analysed from the perspective of perception, whereby listeners' judgements, evaluations and ratings are obtained in perception tests. Additionally, it is well established that poor prosody affects intelligibility and comprehensibility in spoken language communication to a degree that is at least comparable with segmental pronunciation errors (Anderson-Hsieh & Koehler, 1988; Brahimi, Boula de Mareuil & Gendrot, 2004; Boula de Mareuil & Vieru-Dimulescu, 2006; Kennedy & Trofimovich, 2008; Major, Fitzmaurice, Bunta & Balasubramanian, 2002; Munro & Derwing, 1995, 1998, 2001; Tajima, Port & Dalby, 1997).

The importance of both segments and prosody in speech production and perception of FA has been addressed in a number of studies:

- perception studies have focused on the investigation of segments and prosody and their contribution to the degree of perceived FA (Anderson-Hsieh, Johnson & Koehler, 1992; Holm, 2008; Jilka, 2000; Magen, 1998; Munro, 1995; Munro & Derwing, 1995; Trofimovich & Baker, 2006, 2007; Vassiere & Boula de Mareuil, 2004), and
- production studies have shown that L2 speakers struggle equally with segmental (i.e. vocalic: Baker & Trofimovich, 2005; Flege, McKay & Meador, 1999; Fox, Flege & Munro, 1995; Ingram & Park, 1997; Sancier & Fowler, 1997; Walley & Flege, 1999, and consonantal: Bradlow, Pisoni, Akahane-Yamada & Tohkura, 1997; Escudero & Boersma, 2004; Trofimovich, Gatbonton & Segalowitz, 2007; Tsukada, Birdsong, Mack, Sung, Bialystok & Flege, 2004) as well as prosodic aspects of speech (Davidson, 2006; Jilka, 2007; Saito, 2006).

Further studies have identified numerous sources and causes for the production and perception of FA in L2 speech:

- different acquisition scenarios, i.e. simultaneous multilingualism/bilingualism (Genesee, 2000; Meisel, 2001, 2004) vs. sequential multilingualism/bilingualism (Lakshmanan, 1994; McLaughlin, 1978),
- the mode of language learning (instructed vs. naturalistic language learning, DeKeyser, 2003; Ellis, 2006, 2007; Isemonger, 2007), and
- other external factors including sociological and cultural background, situational, contextual and procedural variables and learner-specific factors (including cognitive, motivational and emotional

factors and personality traits, Bayley & Langman, 2004; Bayley & Regan, 2004; Dewaele, 2004; Dörnyei, 2009, Flege, Birdsong, Bialystok, Mack, Sung & Tsukada, 2006; Major, 2004; see Piske, MacKay & Flege, 2001, for an overview).

Several proposals have been put forward over the years to model acquisition processes. Increasingly these proposals adopt a multidimensional approach integrating theoretical linguistics, sociology, and psychology to provide a coherent account of how language is learned (for reviews see Gass, 1988; Thomas, 2005; also Foulkes & Docherty, 2006).

### ***Bidirectional transfer and interference between L1 and L2***

One of the key issues debated with regard to the production and perception of FA in L2 speech is transfer and interference phenomena between L1 and L2 (Atterer & Ladd, 2004; Brière, 1966; Flege, 1981; Wode, 1978, 1992). Transfer and interference phenomena are distinguished in the present paper on the basis of Kellerman and Sharwood Smith's (1986) differentiation between incorporation of elements from one language into the other and overall cross-linguistic influence. Most of the evidence for interference has been provided for the influence of L1 characteristics on L2 showing that they can facilitate and hinder the acquisition process (Gass & Selinker, 1992; Odlin, 1989).

However, current research has begun to investigate cross-linguistic interference going beyond the influence of L1 characteristics on L2. And although L2 influence on L1 has by now been attested for almost all areas of linguistic competence (morphosyntax, pragmatics and rhetoric, the lexicon and semantics; for an overview see Pavlenko, 2000), it seems to be most comprehensively documented for the phonological level. Studies indicating that late acquired L2 phonology does have effects on L1 exist since the 1970s (Fischer-Jørgensen, 1968; Williams, 1979) but are predominantly concerned with the voice onset time of stops (VOT; an acoustic cue known to differ between various languages; Flege, 1987, 2002; Flege & Eefting, 1987; Major, 1992, 1993; Williams, 1980).

Sancier and Fowler (1997) provide acoustic evidence that VOT in the speech produced by a bilingual speaker of American English and Brazilian Portuguese differs according to exposure to the respective language environment. These observations are explained on the basis of "gestural drift" (a change in the articulatory settings (or gesture) associated with a specific sound segment as a result of ambient language imitation, Kuhl & Meltzoff, 1996). Adopting the notion of cross-linguistic category correspondence as proposed by Flege (1987), the changes found in VOT of the speaker are interpreted as the result of a perception-action relationship (Case,

Tuller, Ding & Kelso, 1995). The implications of these findings are twofold: firstly that the organisation of an L1 system can be affected after L1 maturation, and secondly that language learning is based on input, suggesting a usage-based approach. A usage-based theory assumes the emergence of regularities in the linguistic system to be the consequence of frequency of occurrence and co-occurrence. This implies that categorization of individual components and their interaction is the result of experience and repeated exposure to actual sensory objects (Silverman, 2011).

Other studies of more or less isolated phenomena such as intonation and allophonic realizations of other phonemes have added further support to the claim that the phonetic/phonological system of L1 can be accessed and altered even after L1 maturation (Andrews, 1999; Mennen, 2004; Trofimovich & Baker, 2006). Mennen (2004), for instance, studied the intonation pattern of prenuclear pitch accents (LH\*) in the speech production data of native Dutch adult learners of L2 Greek. Although the overall shape of the intonation pattern has been shown to be comparable in that both languages have rising intonation, the timing of f<sub>0</sub> peaks is earlier in Dutch than in Greek; and in Dutch, but not in Greek, phonological vowel length influences peak timing. Since the peak timing of these rising patterns is not categorical and therefore not phonologically contrastive, the cross-linguistic difference has been attributed to phonetic implementation of phonologically identical tonal events. A comparison between native speakers of Dutch and Greek showed that Dutch native speakers transferred f<sub>0</sub> timing of rising pitch accents from their native language into their L2, realising early peaks (within the vowel of the accented syllable) compared to late peaks (realised within the unaccented vowel of the following syllable). This is consistent with previous studies showing influence of L1 intonation patterns on L2. Novel in Mennen's analysis was the finding of L2 influence on L1 intonation patterns, an effect that prior to her study had only been documented for the segmental properties of speech production. She found that Dutch L2 speakers of Greek tended to neutralise the difference in f<sub>0</sub> timing associated with phonological vowel length, realising both short and long vowels with the peak alignments typically found in the "long vowel set" produced by native speakers of Dutch. Willems (1982) found that size and direction of pitch movements in native English speakers acquiring Dutch differed systematically from monolingual L1 English productions, and the differences were attributed to the influence of Dutch intonation patterns.

### ***Implications for SLA models of speech production and perception***

The nature of interference between L1 and L2 has not been formally and comprehensively accounted for in

current models of speech learning which aim at explaining which parts of an L1 influence L2 (and vice versa). In relation to speech production and perception, necessary adaptations of current models of SLA (Kuhl, 1991: Native Language Magnet (NLM), Best, 1995; Best, McRoberts & Goodell, 2001: Perceptual Assimilation Model (PAM), and Flege, 1995; Flege et al., 2006: Speech Learning Model (SLM)) primarily focus on perceptual categories and phonological contrast based on phonetic similarities and differences between the segmental levels. However, it is clear that these models also need to incorporate prosodic characteristics and their interaction with the segmental level. Previous studies, one conducted by Trofimovich and Baker (2006), suggest that L2 segmental learning and L2 prosodic learning seem to be similar in that they appear to be gradual and their development is a result of experience and exposure.

The idea of experience and exposure as factors in L2 production has been addressed in a number of studies of segmental characteristics (Fowler, Sramko, Ostry, Rowland & Halle, 2008; Major, 1992; Sancier & Fowler, 1997). Crucially though, Trofimovich and Baker (2006) state that the learning of L2 prosody differs with respect to individual prosodic characteristics, i.e. speech melody vs. speech fluency. A tentative conclusion drawn from their results is that those characteristics that are perceptually salient or relatively distinct in native and target language are more easily acquired than those that are perceptually similar. These results can be interpreted in favour of the SLM and its notion of “equivalence classification”. Trofimovich and Baker’s observation of experience effects in prosodic characteristics lends support to a core assumption of perceptual models of SLA, namely that input and use guide the L2 acquisition process. The next step towards an integrated model of speech production and perception in SLA is the attempt to untangle the interdependencies between cross-linguistic prosodic and segmental similarities and differences, ideally under the consideration of other linguistic and extra-linguistic factors mentioned above. This will, however, leave the field with a multitude of questions to be answered. The current study aims at contributing to this enterprise.

### **1.2 Regional variation and its impact on SLA**

The study of language variation has a long-standing tradition in sociolinguistics (Labov, 1972). One of the best understood and documented causes of variability is differences between dialects (or regional varieties; these terms are interchangeably used throughout the paper referring to regionally distinctive varieties of a language, identifiable by a particular set of lexical, grammatical and phonological characteristics; Chambers & Trudgill, 1998). Sociophonetics with an emphasis on regional variation has provided evidence that sound systems

and their organisation can differ between dialects as much as between languages, including regional prosodic characteristics (Gooskens, 1997; van Leyden & van Heuven, 2003), and it seems rather surprising that the influence of regional aspects has been largely neglected in the context of SLA.

#### ***Regional variation in L1 and its influence on L2***

The influence of regional L1 characteristics on L2 speech has, for instance, been addressed by Atterer and Ladd (2004). They found that cross-varietal differences between Northern and Southern German in pitch alignment are carried over into pitch accent realizations in L2 English, in that Southern German speakers align rises (L\*H) in prenuclear accented syllables later compared to peak realization in Northern German varieties. Addressing segmental characteristics, Teasdale (1997) explains the substitution of /θ/ with [t] by speakers of Québécois French and its substitution with [s] by European French speakers on the basis of differences in the place of articulation for /s/ between the two groups of speakers, whereby the former produce the voiceless fricative in alveolar and the latter in dental position.

#### ***Regional variation in L2 and its influence on L2***

A very limited number of studies have shown that differences and similarities in the organization of sound systems between varieties of an L2 influence accurate discrimination and identification of L2 sounds. Escudero and Boersma (2004) as well as Baker & Smith (2010) studied differences in the vocalic systems of varieties of a target language. Escudero and Boersma (2004) provide evidence that production and perception of /i–i/ contrast acquired by native speakers of Spanish differ as a result of exposure to either Scottish English or Southern British English. Baker & Smith (2010) found that native North American English speakers acquire the French /i–y–u/ vowel contrast differently as a result of exposure to Québécois French or European French. Additionally, they provide evidence that additional acoustic cues that are not inherently included in the vowel contrast itself, but arise from the vowel’s phonetic environment (i.e. assibilation or affrication of alveolar plosives before high vowels) influence accuracy in production and perception.

#### ***Regional variation and its influence on L2 perception***

Regional variation has also been addressed in the context of foreign accent perception. For example, a study by Ikeno and Hansen (2006) examines the influence of listener’s accent background on accent perception and comprehensibility. British English varieties spoken in Cardiff, Cambridge and Belfast were presented to native speakers of British and American English as well as non-native listeners from varying linguistic

backgrounds. The results showed that stimulus context length, listener's national and regional background and comprehensibility of speech samples influence accuracy in accent detection and classification. Two findings are of particular interest for the present study. Firstly, Belfast English was perceived as non-native English by American listeners and non-native listeners with varying linguistic background whereas British English listeners accurately identified the Belfast accent in nearly 100% of the stimuli presented. Secondly, the group of British English listeners confused Cardiff and Cambridge accents significantly more frequently than the Belfast English accent with the Cardiff and Cambridge accents. This suggests that Belfast English provides phonetic cues that are perceptually more different compared to those of the other two varieties investigated and which may also possibly be closer to the British standard pronunciation.

Another study conducted by van Bezooijen and Gooskens (1999) used manipulated speech samples that suspended portions of the acoustic signal to investigate the role of prosody, pronunciation and other linguistic information used by Dutch and British English listeners to detect regional varieties of their respective native language. In addition to integral samples (that contained all linguistic information), low-pass filtered speech samples with only prosodic information and monotonised speech where intonation had been removed were presented to native speakers of Dutch and British English. The listeners were asked to identify the speakers' regional origin. What seems particularly interesting in the context of the present paper is a category added in the English part of the experiment pertaining to regional varieties of British English. The category was added since prosody in general and intonation in particular has been studied more extensively in British English compared to Dutch and consisted of speech samples identified by English and Dutch linguists specialized in intonation and English varieties. Samples were selected only on the basis of perceivable prosodic characteristics specific to regional varieties of British English as opposed to all other samples that were randomly selected on the basis of the speakers' origin. When presented to native speakers of Dutch and British English, both groups of listeners performed comparably well in the identification of regional origin based on verbal and pronunciation cues. English listeners however performed significantly better in the identification of speech samples on the basis of prosodic information only, which was found to be nearly impossible for the Dutch listeners. Significant differences were found between specifically and randomly selected stimuli elicited from the Belfast English data set.

Crucially, both studies provide evidence that (i) Belfast English possesses identifiable regional prosodic characteristics and (ii) that identification of regional varieties critically depends on the listener's linguistic

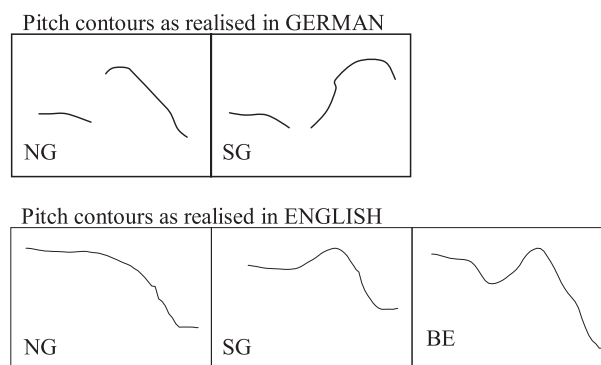


Figure 1. Illustration of German nuclear pitch accents of *im September* extracted from the present corpus produced by native speakers of NG and SG (top row) and English nuclear pitch accents of *MANgo* produced by native speakers of SG, NG and BE.

background, i.e. his/her experience, exposure and environment.

### 1.3 Prosodic characteristics of Swiss German, Northern Standard German and Belfast English

In order to contribute to our understanding of the nature of interference between L1 and L2, we investigate cross-linguistic similarities and difference between regional varieties (Northern Standard German, Swiss German spoken in Bern and English spoken in Belfast) and their influence on the production and perception of L1 and L2. Prosodic characteristics have been shown to differ systematically between languages and also regional varieties. In particular, hypotheses about typological characteristics of languages and varieties have been established on the basis of such systematic differences in the realization of pitch accents. Previous analyses of pitch accent realizations in varieties of German and English indicate that Swiss German, particularly Bernese German, and Belfast English share specific characteristics of nuclear pitch accent realizations. Bernese German (Fitzpatrick-Cole, 1999) as well as Belfast English (Lowry, 2002) have been shown to feature rising or low-targeted pitch patterns in nuclear accents of declarative utterances while most other varieties of German and English were found to have falling or high-targeted pitch patterns in declaratives (Atterer & Ladd, 2004; Grabe 2002, 2004; Ulbrich, 2004, 2005). Figure 1 provides an illustration of typical pitch contours realised by Swiss German (SG), Northern Standard German (NG) and Belfast English (BE) speakers in identical utterances extracted from the present corpus.

In conclusion, the current study draws on several relevant issues addressed in previous research: (i) L2 has been shown to influence matured L1, (ii) regional variation in both native and target language influences

language production of a L2, (iii) regional variation influences the perception of foreign accentedness, and (iv) these characteristics have been documented for both the segmental and the prosodic level of speech. It seems therefore reasonable to assume that regional characteristics in L1 and L2 also impact on L1. The aim of the present study is thus to investigate to what extent cross-linguistic and cross-varietal similarities and differences between regional varieties of L1 German and L1 and L2 English affect L2 and L1 production as well as L2 perception.

In order to do so, the paper reports the analysis of pitch accent realisations of native German L2 learners of English with and without exposure to Belfast English to show that L2 speakers are able to acquire prosodic characteristics of a target language. Furthermore, the comparison of L2 Belfast English produced by native Swiss German and Northern Standard German speakers with exposure to Belfast English will elucidate the impact of cross-varietal differences in pitch accent realisations in L2. Foreign accent ratings of L2 Belfast English stimuli obtained from native speakers of Belfast English will explore if these differences are perceivable by native speakers. The comparison of German utterances produced by native German L2 learners of English with and without exposure to Belfast English will show whether or not specific characteristics of pitch realisation of L2 Belfast English have an impact on L1. The results will be interpreted in adaptation of a usage-based approach.

## 2. Experiment 1: Cross-varietal and cross-language differences in the production of pitch accents

The experiment described in this section is designed to provide evidence that cross-linguistic and cross-varietal differences are reflected in L2 speech production.

### 2.1 Method

The experiment is based on recordings of native speakers of Belfast English and three groups of native German speakers of L2 English: Bernese German speakers, Northern Standard German speakers with exposure to Belfast English, and Northern Standard German speakers without exposure to Belfast English. The analysis of speech production data is based on (i) L1 German produced by the three groups of native German speakers, and (ii) a comparison of their L2 English with the native Belfast English speakers. The differentiations between regional varieties in L1 and L2 provide the basis for cross-linguistic and cross-varietal comparisons. The perception experiment will shed light onto the question of whether cross-language similarities of nuclear pitch patterns in a variety of L1 German influence FA perception in L2 English by native speakers of Belfast English.

### 2.2 Speech materials

The present study is based on recordings of four groups of speakers:

- native speakers of Belfast English (BE)
- native Swiss German speakers of the Bernese dialect with no previous exposure to Belfast English (SG)
- native German speakers of Northern Standard German who have lived in Belfast for a minimum of three years (NG+ex)
- native German speakers of Northern Standard German with no previous exposure to Belfast English who have never lived in an English-speaking environment and acquired English as a foreign language at school in Germany (NG-ex).

NG+ex and NG-ex speakers were from Hanover and the surrounding area, producing a variety of German that is the most similar to the Northern German Standard variety (e.g. Clyne, 1995). We recorded a total of 40 subjects; 20 female and 20 male speakers (five per speaker group). The recordings were carried out in a quiet room at the University of Ulster using a Sennheiser ME64 directional condenser microphone (cardioid, frequency response 40 Hz – 20,000 Hz,  $\pm 2.5$  dB) with a sampling rate of 22,050 Hz directly onto a Toshiba notebook computer for processing and analysis in PRAAT (Boersma & Weenink, 2005). Statistical analyses were carried out in SPSS.

The speakers were aged between 20 and 52 years, with the following average ages: NG-ex\_female: 22.8; SG\_female: 23.8; NG+ex\_female: 34.4; BE\_female: 29.2; NG-ex\_male: 23.2; SG\_male: 22.8; NG+ex\_male: 36.6 and BE\_male: 34. SG and NG-ex subjects were students or exchange students at the University of Ulster or Queen's University in Belfast. Some of the Swiss German speakers were visiting friends. Neither SG nor NG-ex speakers had been exposed to the variety of English spoken in Belfast prior to the experiment. Also, they had not spent more than three weeks at any one time in an English-speaking environment. The group of NG+ex speakers was selected under the criteria that they have not been exposed to any other English variety prior to their arrival in Belfast and that they have experienced long-term exposure to the Belfast English variety, which explains that their average age was higher compared to the other groups of speakers. Therefore age of learning could not be considered in the current study.

We did not carry out any assessment of fluency prior to the study although it has been shown to influence L2 production and perception (Gut, 2009). This was not deemed to be necessary since the acoustic manipulation procedure applied to create the stimuli for the perception test had to be based on utterances of comparable length which did not include pausing or hesitations or differences

in accent placement. No subjects received payment for their participation.

The subjects were given a reading task. This was considered essential since the stimuli creation for the perception task had to be based on directly comparable speech data. The task was carried out by all groups of speakers in English and additionally in German by the native speakers of German. The subjects were instructed to read a short text. Embedded in the text were nine declarative utterances, each containing one of nine target words. The target words were placed at the end of a short paragraph in order to avoid pitch contours indicating continuation, which have been shown to differ considerably between regional varieties (Gilles, 2005) and to distinguish them from pitch patterns indicating termination (Grice & Bauman, 2007). The targets – five bi-syllabic and four tri-syllabic words – are cross-linguistically comparable with respect to their segmental content and stress placement. They have been placed in nuclear accent position of short, broad focus utterances of comparable length. The distinction between bi-syllabic and tri-syllabic target words was introduced to the corpus since impressionistic observation indicated differences of pitch accent realisation on the basis of syllable number. An example for each type of target word is given below.

Example 1: German: Das Haus hatte einen kleinen **Garten**.

English: The house had a little **garden**.

Example 2: German: Er übergab ihr das **Telefon**.

English: He handed her the **telephone**.

A total of 630 target sentences were extracted from the recorded readings for further analysis, nine target sentences per speaker in English, five containing bi-syllabic and four containing tri-syllabic words (90 per group of speakers; 360 English sentences overall) and nine target sentences per speaker in German produced only by the three groups of native German speakers, again five containing bi-syllabic and four containing tri-syllabic words (90 per group of speakers; 270 German sentences overall).

### 2.3 Procedures for the analysis of speech production data

The entire corpus was segmented and phonetically transcribed by a phonetician using the IPA based on perception and visual clues provided in spectrograms. The nuclear pitch accents were intonationally labelled using the rhythmic and the phonetic tier in an adaptation of the IViE system (Grabe, 2004). The prosodic annotation of nuclear pitch patterns was carried out by two to four trained annotators, blind to conditions. Two of the annotators were German, fluent in English and living in

Belfast, two annotators were native speakers of English. Nuclear pitch accents were phonetically labelled to capture the general directionality of pitch patterns, the actual realisation of high or low pitch targets within the accented syllable (i.e. the alignment of the f<sub>0</sub>-contour within the accented syllable), and the pitch contour following the target. The following realisations were distinguished:

High-target pitch accents	(H)
Low-target pitch accents	(L)
Low target rising pitch accents	(LH)
High-target falling pitch accents	(HL)
Low-target rising-falling pitch accents	(LHL)

In order to test for reliability of annotations across the annotators we compared the results of the labelling. 78% of the scores were consistent across annotators. The majority of variability (17%) was found in annotations of rising patters (LH was frequently labelled as LHL and vice versa). In such cases, the annotators revisited the individual pitch patterns, and decisions were made collaboratively after visual examination of the f<sub>0</sub> contour. Since we were not interested in a phonological typology of high or low targeting pitch accents and boundary tone realisations in the sense of the autosegmental-metrical approach (Pierrehumbert, 1980), the prosodic annotation was carried out to provide confirmation of cross-varietal and cross-language differences in the present corpus that have previously been found and described in the literature (see Section 1 above). A PRAAT script was employed to measure f<sub>0</sub> (st) and duration (ms) at two (potentially three) points within each voiced portion of each syllable transcribed: at the beginning, the end and at potentially appearing f<sub>0</sub> turning points in the pitch contour. The measurements were manually inspected and values were subsequently returned into separate tiers for further analyses of f<sub>0</sub> movement and alignment (see Section 3 below).

### 2.4 Results

The following section reports the findings of cross-linguistic and cross-varietal comparisons of the realisation of nuclear pitch patterns produced by the four groups of speakers (BE, SG, NG+ex, and NG-ex) in the two languages (German and English).

The analysis of the recorded speech material provides evidence for both cross-varietal and cross-linguistic comparisons in the realisation of nuclear pitch patterns. The first subsection below reports the findings of cross-varietal differences in the German speech production comparing nuclear pitch accent realisations across the three groups of native German speakers (SG, NG+ex and NG-ex). Cross-linguistic comparisons revealed

Table 1. Average % of pitch patterns in bi-syllabic (BI) and tri-syllabic (TRI) German target words produced by speakers of NG-ex, NG+ex and SG.

	H		HL		L		LH		LHL		TOTAL H		TOTAL L	
	BI	TRI	BI	TRI	BI	TRI	BI	TRI	BI	TRI	BI	TRI	BI	TRI
NG-ex	6	7	74	85	14	1	12	2	0	5	80	92	20	8
NG+ex	0	4	56	63	6	0	16	11	22	22	56	67	44	33
SG	18	6	9	18	12	6	7	7	54	63	27	24	73	76

differences between BE, SG, NG+ex and NG-ex in the English production data as detailed in the next subsection. The results of a comparison between L1 German and L2 English across the three groups of native German speakers are detailed in the final part of Section 2.4.

#### Comparison of SG, NG + ex and NG-ex in German data

A multivariate ANOVA with *Origin* (with three levels: SG, NG+ex and NG-ex) and *Gender* (with two levels: female and male) as between subject factor and *Pitch Target* as dependent variables (BI\_H, BI\_HL, BI\_L, BI\_LH, BI\_LHL, TRI\_H, TRI\_HL, TRI\_L, TRI\_LH and TRI\_LHL) revealed significant cross-varietal differences between the three groups of native German speakers with a considerable effect size [Pillai's Trace:  $F_{(18;40)} = 12.030$ ,  $p < .0001$ ,  $\eta^2 = .844$ ]. To reduce type 1 errors, Bonferroni adjustments ( $p < .0167$ ) for multiple pair-wise post-hoc comparisons were applied. The difference between female and male speakers was not significant, consequently no interactions were found between *Gender* and *Origin*. The combined results are presented in Table 1. Percentages are provided for each of the annotated pitch patterns (H, HL, L, LH, and LHL). The two rightmost columns give the overall percentage for H target pitch patterns (pooling H and HL) and for L target pitch patterns (pooling L, LH, and LHL). Multiple pair-wise comparisons showed that differences were not found in the realisations of all pitch targets. The three groups of speakers did not differ significantly in the realisation of un-pooled H and L pitch patterns in both bi- and tri-syllabic target words. Equally no differences were found in the realisation of rising pitch accents (LH) in tri-syllabic words. Note, however, that these pitch targets were only observed in 13% of all target words (36 realisations in 270 examined pitch accents).

Considering the distinction between bi- and tri-syllabic targets (150 and 120 realisations respectively), the percentage of observations in each of these tonal categories is invariably below 5% with the exception of L pitch accents in bi-syllabic words found in 9% of the realisations (14 cases). In tri-syllabic words, three targets (2%) were produced with a low (L) pitch target. Rising pitch patterns (LH) were found in seven tri-syllabic words (6%). Speakers realised high targets (H) only in

eight bi-syllabic (5%) and five tri-syllabic words (4%). In the realisation of the latter (H in tri-syllabic words), the difference between all three groups of speakers was insignificant [ $F_{(2;26)} = 2.61$ ,  $p < .331$ ]. In the realisation of tri-syllabic target words with a low target (L) and bi-syllabic target words with a high target (H), only NG-ex and NG+ex did not differ significantly. SG speakers produced significantly more low targets (L) in tri-syllabic words [ $F_{(2;26)} = 11.8$ ,  $p < .0001$ ] and H targets in bi-syllabic words [ $F_{(2;27)} = 15.5$ ,  $p < .0001$ ].

Low targets (L) in bi-syllabic words, however, were produced less often by NG+ex speakers compared to both NG-ex and SG speakers, and were produced most often by NG-ex speakers [ $F_{(2;27)} = 7.2$ ,  $p < .027$ ].

Significant differences between all three groups of speakers (SG, NG+ex and NG-ex) were found in the realisation of HL, LH, and LHL in bi-syllabic words [ $F_{(2;27)} = 167.035$ ,  $p < .0001$ ;  $F_{(2;27)} = 69.4$ ,  $p < .0001$ ;  $F_{(2;27)} = 80.3$ ,  $p < .0001$ , respectively] and the realisation of HL and LHL in tri-syllabic target words [ $F_{(2;26)} = 156.2$ ,  $p < .0001$ ;  $F_{(2;26)} = 147.2$ ,  $p < .0001$ , respectively]. Overall, rising-falling pitch patterns (LHL) were produced in the majority by NG+ex and SG speakers and differed significantly in both bi-syllabic (21%) and tri-syllabic targets (29%). Falling pitch accent realisations (HL) were found in 45% of bi-syllabic and 56% of tri-syllabic target words. Rising pitch accents (LH) were produced in 19% of bi-syllabic target words.

SG speakers realised overall more low target pitch patterns compared to NG-ex and NG+ex but produced in the majority of cases rising-falling (LHL) pitch accents. NG-ex speakers on the other hand produced mainly falling (HL) pitch accents. NG+ex speakers produced significantly more rising-falling pitch accents (LHL) compared to NG-ex speakers but still produced predominantly falling pitch patterns in nuclear accent positions (HL). The difference between the three groups of native speakers of German (SG, NG+ex and NG-ex) is most visible in a direct comparison of average percentages of pitch accent realisations with high and low pitch targets as illustrated in the two rightmost columns of Table 1. NG-ex speakers produced overall 86% of the target words with a high target and only 14% with a low target. Nearly the exact opposite is the case in realisations of SG speakers.



Table 2. Average % of pitch patterns in bi-syllabic (BI) and tri-syllabic (TRI) English target words produced by speakers of NG-ex, NG+ex, SG and BE.

	H		HL		L		LH		LHL		TOTAL H		TOTAL L	
	BI	TRI	BI	TRI	BI	TRI	BI	TRI	BI	TRI	BI	TRI	BI	TRI
NG-ex	4	0	90	86	0	8	3	0	3	6	94	86	6	14
NG+ex	0	0	42	31	8	4	28	14	28	52	40	32	60	68
SG	0	0	43	37	0	0	32	8	26	58	42	38	58	62
BE	0	0	27	13	0	0	26	32	48	54	26	14	74	86

Here 75% of the targets were realised with a low target and only 25% with a high target.

The NG+ex speakers realised 62% of the target words with a high target but 38% with a low target. As for the NG-ex and the SG groups, these results confirm expectations derived from previous investigations of pitch accent realisations in Northern Standard German and Bernese German (Fitzpatrick-Cole, 1999; Ulbrich 2004, 2005).

#### Comparison of BE, SG, NG+ex and NG-ex in English recordings

The same statistical test was applied to the English production data as was applied previously to the German speech material, i.e. a multivariate ANOVA with *Origin* (this time with four levels: BE, SG, NG+ex and NG-ex; including the native speakers of Belfast English) and *Gender* (male vs. female) as between subject factors and *Pitch Target* as the dependent variables (BI\_H, BI\_HL, BI\_L, BI\_LH, BI\_LHL, TRI\_H, TRI\_HL, TRI\_L, TRI\_LH and TRI\_LHL). Again to control for error type I, Bonferroni adjustments ( $p < .0125$ ) were applied for multiple pair-wise comparisons. Between the four groups of speakers, a significant main effect provides evidence for cross-varietal differences [Pillai's Trace:  $F_{(27;90)} = 9.73$ ,  $p < .0001$ ,  $\eta^2 = .745$ ]. No significant differences were found between female and male speakers, hence no interactions between *Origin* and *Gender*. The results are presented in Table 2.

Similar to the German part of the corpus, the pitch accent realisations most infrequently found in the English data did not differ significantly between the four groups of speakers; this includes H and L realisations in bi- and tri-syllabic target words. H was only realised twice in bi-syllabic targets (by one NG-ex male speaker) whereas L was realised two and four times in bi- and tri-syllabic words by individual NG-ex and NG+ex speakers and therefore concerns a negligible 2% of all realisations.

The four groups of speakers (NG-ex, SG, NG+ex and BE) differed significantly in the realisation of HL, LH, and LHL in bi-syllabic [ $F_{(3;36)} = 65.24$ ,  $p < .0001$ ;  $F_{(3;36)} = 42.1$ ,  $p < .0001$ ;  $F_{(3;36)} = 54.3$ ,  $p < .0001$ , respectively]

and tri-syllabic target words [ $F_{(3;35)} = 91.7$ ,  $p < .0001$ ;  $F_{(3;35)} = 32.3$ ,  $p < .0001$ ;  $F_{(3;35)} = 33.6$ ,  $p < .0001$ , respectively]. Overall, in the English data, falling pitch patterns (HL) were realised most frequently (52% in bi-syllabic and 43% in tri-syllabic target words). 67% of the target words were realised with rising-falling (LHL) pitch accents, i.e. 26% in bi-syllabic and 41% in tri-syllabic target words. Rising pitch patterns (LH) were found in 24% of bi-syllabic and 11% of tri-syllabic target words.

As before, the differences between the four groups of speakers are most obvious in the direct comparison of H and L targets (illustrated in the two rightmost columns of Table 2). NG-ex produced bi- and tri-syllabic words mostly (>85%) with a falling (HL) pitch pattern, thus significantly more often compared to the other three groups of speakers. BE speakers by contrast produced less than a third of bi- and tri-syllabic words with a falling accent (27% and 13%, respectively), significantly less often than both NG+ex and SG speakers (NG+ex: 42% and SG: 43% in bi-syllabic words; NG+ex: 31% and SG: 37% in tri-syllabic words). SG and NG+ex speakers did not differ significantly. In tri-syllabic target words, BE speakers produced significantly more rising pitch patterns (LH) compared to all three other groups of speakers. However, NG+ex speakers produced rising pitch accents significantly more often than both SG and NG-ex speakers.

The analysis of LH in bi-syllabic and LHL in tri-syllabic target words revealed similar results. The patterns were only realised in 3% and 6% by NG-ex speakers, significantly less frequently compared to SG, NG+ex and BE groups. Additionally, BE speakers differed significantly from NG+ex and SG speakers in the realisation of rising-falling (LHL) pitch patterns in tri-syllabic words, producing rising-falling pitch patterns significantly more often than SG and NG+ex. Although BE speakers realised predominantly rising and rising-falling pitch accents, approximately one third of the pitch realisations were produced with falling pitch accents (HL). This deviates somewhat from findings presented by Grabe (2004) who analysed pitch realisations in a number of regional varieties of English. She found 83.3% of declaratives produced by Belfast speakers of English to be

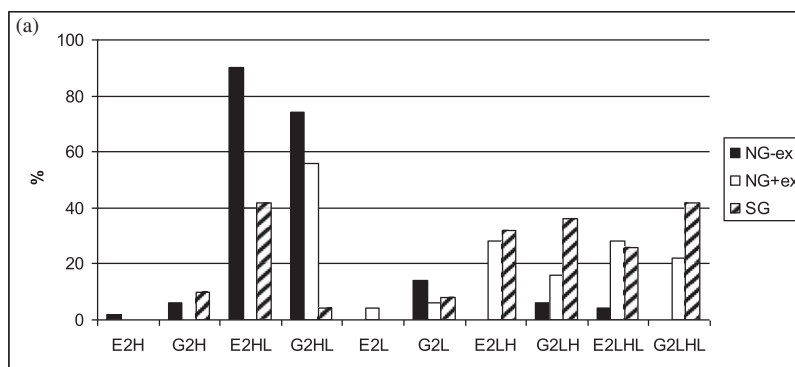


Figure 2a. Comparison of NG-ex, NG+ex and SG speakers' pitch accent realisations in bi-syllabic words (indicated by 2) produced in English (E) and German (NG-ex) data.

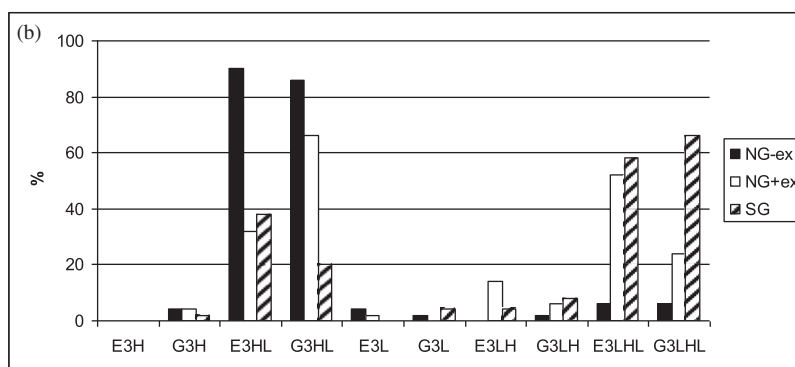


Figure 2b. Comparison of NG-ex, NG+ex and SG speakers' pitch accent realisations in tri-syllabic words (indicated by 3) produced in English (E) and German (G) data.

realised with rising intonation. The deviation might be due to differences in the speaking style since the analysis in the present study was based solely on read speech, which presumably is more formal than spontaneous speech. For instance, Lowry (2002) found differences in the number of rising pitch patterns produced by native speakers of Belfast English depending on speaking style. Her results show 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, a variety that features predominantly falling pitch accents in nuclear accents of declaratives).

#### **Comparison of English and German recordings in SG, NG+ex and NG-ex**

A comparison between the German and the English data produced by the three groups of German native speakers (NG-ex, NG+ex and SG) revealed some interesting results. Both German and English data of the three groups were submitted to a mixed ANOVA again with Bonferroni adjustments ( $p < .0046$ ) for multiple pair-wise comparison with *Pitch Accent* (10 levels: 2\_H, 2\_HL, 2\_L, 2\_LH, 2\_LHL, 3\_H, 3\_HL, 3\_L, 3\_LH, 3\_LHL) and *Language* (2 levels: German and English) as within

subject factors and *Origin* as between subject factor (NG-ex, NG+ex and SG). As the sphericity assumption was violated, Greenhouse-Geisser corrections were applied and are reported. The analysis revealed a significant main effect [ $F_{(9,4:896)} = 9.8, p < .0005, \eta^2 = .267$ ] for *Pitch Accent* depending on *Language*, but in addition we found a significant interaction with the factor *Origin* with a considerably larger effect size [ $F_{(18,9:783)} = 31.5; p < .0005, \eta^2 = .700$ ]. The results of the cross-linguistic comparison are illustrated in Figures 2a and 2b.

The interaction is due to the fact that across all possible realisations of German and English bi- and tri-syllabic target words, NG-ex speakers most often produced falling pitch patterns (HL) whereas NG+ex and SG speaker's pitch accent realisations differed significantly depending on the language. NG+ex speakers produced considerably more rising (LH) and rising-falling (LHL) pitch accents in both bi-syllabic and tri-syllabic target words in the English data compared to the German data, indicating that they have acquired regionally marked pitch patterns. The results confirm previous findings of FA ratings obtained from native speakers of Belfast English (Ulbrich, 2008:), where NG+ex learners of Belfast English received significantly lower FA scores compared

to NG+ex speakers of English with no exposure to the Belfast English variety. Previous and present findings strengthen our interpretation that NG+ex speakers acquire characteristics of a regional variety which in turn affects the perception of FA. However, it is worth pointing out again that NG+ex speakers produce rising and rising-falling pitch accents significantly more often than NG-ex speakers also in L1 German. These findings suggest that L2 characteristics influence L1 since both NG+ex and NG-ex speakers, i.e. native German speakers of L2 English with and without exposure to the English variety spoken in Belfast, have the same regional L1 background. All German speakers were speakers of a Northern Standard variety of German which features a falling default nuclear pitch accent in declaratives (see Section 1). Hence, the realisation of more LH and LHL patterns which are associated with a default accent of Belfast English in declaratives may be the result of L2 interference with L1.

The SG data also show some interesting results. SG speakers realised considerably more falling (HL) pitch patterns in the English data compared to the German data. The difference was found to be significant in both bi- and tri-syllabic target words. Since none of the recorded SG speakers had any previous experience with the regional variety of English spoken in Belfast or had lived for longer than three weeks in an English-speaking environment, these findings could be interpreted as a result of classroom instruction. The more frequent realisation of falling accents (HL) is accompanied with a decreasing number of rising-falling pitch patterns (LHL) in the English data compared to the German.

### 3. Experiment II: Foreign accent perception in L2 English

The perception experiment was based on English utterances only. To investigate the cross-language influence as a result of (i) cross-varietal prosodic differences in L1 German, and (ii) exposure to an L2 variety, the experimental design had to allow for a threefold direct comparison between SG vs. NG-ex; NG+ex vs. NG-ex and SG vs. NG+ex speakers as well as their individual comparison to the group of native speakers of Belfast English. In order to facilitate the investigation of cross-varietal difference between L2 English produced by the three groups of German speakers, two native speakers of Belfast English were chosen from the present corpus – one female and one male – as the basis for acoustic manipulations. To examine if exposure to L2 Belfast English (for NG+ex) and cross-language similarities in pitch accent patterns between Belfast English and Swiss German (for SG) leads to comparable pitch accent realisation in L1 German we also chose one female and one male NG-ex speaker as a basis for acoustic

manipulations. Both BE and NG-ex speakers were chosen on the basis of their perceptually most uncharacteristic voice quality.

#### 3.1 Stimuli preparation

Prior to the stimuli creation, syllables of utterances containing the target words were automatically segmented on the basis of orthographic transcription using the PRAAT algorithm (Boersma & Weenink, 2005). The segmentation has been checked manually by visually inspecting the original waveform and the associated spectrogram. The creation of stimuli for the perception experiment employed a procedure known as prosodic transplantation. This method extracts and processes pitch and duration parameters in application of a TD-PSOLA algorithm (Time Domain Pitch Synchronous Overlap and Add, Moulines & Charpentier, 1990), while energy is normalised. The method has previously been employed in the study of FA perception (Boula de Mareuil & Vieru-Dimulescu, 2006; Brahimi et al., 2004; Ulbrich, 2008;). The algorithm defines  $f_0$  at three points in the periodic portion of each syllable, i.e. at the beginning, the midpoint and at the end. These are connected so that  $f_0$  of each syllable is associated with two linear pitch movements with  $f_0$  of unvoiced segments set to zero. The initial  $f_0$  value of each syllable is connected to the final  $f_0$  value of the preceding syllable.

The extracted prosodic parameters of one voice are then crossed with the segments of another voice. We crossed the segmental level with all prosodic levels ( $4 \times 4$ ) so that we had a total of 16 conditions for our perception experiment as detailed in Table 3.

For the creation of stimuli, we initially crossed segments (indicated throughout the paper by subscript  $s$ ) of selected NG-ex and BE speakers with the prosodic level (indicated throughout the paper by subscript  $p$ ) of SG and NG+ex speakers (female and male speakers respectively) resulting in the conditions:  $BE_sSG_p$ ,  $BE_sNG+ex_p$ ,  $NG-ex_sSG_p$ , and  $NG-ex_sNG+ex_p$ .

Subsequently, we crossed their prosodic level with the segments of the SG and NG+ex speakers to create  $SG_sBE_p$ ,  $SG_sNG-ex_p$ ,  $NG+ex_sBE_p$ , and  $NG+ex_sNG-ex_p$ . The acoustic modification of acoustic parameters such as  $f_0$  or duration creates deteriorations (audible artefacts) that are not found in natural voice qualities. Therefore, we crossed all 40 speaker's original segments with their original prosodic level to create "original" stimuli ( $BE_sBE_p$ ), ( $SG_sSG_p$ ),  $NG+ex_sNG+ex_p$ , and  $NG-ex_sNG-ex_p$  with comparable sound quality of acoustically modified speech. We crossed all NG-ex speakers with one female and one male BE speaker to create  $BE_sNG-ex_p$  and  $NG-ex_sBE_p$  allowing for comparison across the three groups of German speakers. Lastly, we crossed

Table 3. *Manipulations of segmental and prosodic level for stimuli creation (BE = Belfast English).*

Manipulation	Segmental level	Prosodic level
BE <sub>S</sub> BE <sub>P</sub>	Native Belfast	Native Belfast
BE <sub>S</sub> SG <sub>P</sub>	Native Belfast	Swiss German
BE <sub>S</sub> NG+ex <sub>P</sub>	Native Belfast	L2 learners of BE
BE <sub>S</sub> NG-ex <sub>P</sub>	Native Belfast	German no exposure
SG <sub>S</sub> BE <sub>P</sub>	Swiss German	Native Belfast
SG <sub>S</sub> SG <sub>P</sub>	Swiss German	Swiss German
SG <sub>S</sub> NG+ex <sub>P</sub>	Swiss German	L2 learners of BE
SG <sub>S</sub> NG-ex <sub>P</sub>	Swiss German	German no exposure
NG+ex <sub>S</sub> BE <sub>P</sub>	L2 learners of BE	Native Belfast
NG+ex <sub>S</sub> SG <sub>P</sub>	L2 learners of BE	Swiss German
NG+ex <sub>S</sub> NG+ex <sub>P</sub>	L2 learners of BE	L2 learners of BE
NG+ex <sub>S</sub> NG-ex <sub>P</sub>	L2 learners of BE	German no exposure
NG-ex <sub>S</sub> BE <sub>P</sub>	German no exposure	Native Belfast
NG-ex <sub>S</sub> SG <sub>P</sub>	German no exposure	Swiss German
NG-ex <sub>S</sub> NG+ex <sub>P</sub>	German no exposure	L2 learners of BE
NG-ex <sub>S</sub> NG-ex <sub>P</sub>	German no exposure	German no exposure

SG speakers with NG+ex speakers, male and female accordingly resulting in SG<sub>S</sub>NG+ex<sub>P</sub> and NG+ex<sub>S</sub>SG<sub>P</sub>.

The following predictions were established:

- BE<sub>S</sub>BE<sub>P</sub> is expected to receive the lowest FA ratings since these stimuli involve only segments and prosody of the listeners' native variety.
- NG-ex<sub>S</sub>NG-ex<sub>P</sub> is expected to receive the highest FA ratings since the NG-ex speakers have not been exposed to Belfast English so that regional characteristics of this variety cannot influence FA perception. In addition, pitch accent patterns between Northern Standard German and Belfast English differ considerably (see Section 1.3). These differences are expected to increase perceived foreign accentedness.
- NG+ex<sub>S</sub>NG+ex<sub>P</sub> is expected to receive FA ratings between BE<sub>S</sub>BE<sub>P</sub> and NG-ex<sub>S</sub>NG-ex<sub>P</sub>. These speakers have been exposed to the Belfast English and the acquisition of its regional characteristics is expected to reduce the perception of FA.
- Due to cross-language similarities between Swiss German spoken in Bern and Belfast English in pitch accent realisations, lower ratings for crossings involving SG segments (SG<sub>S</sub>) and prosody (SG<sub>P</sub>) compared to those involving NG-ex are expected.
- Furthermore, similar ratings for BE<sub>S</sub>SG<sub>P</sub> vs. BE<sub>S</sub>NG+ex<sub>P</sub>, and NG-ex<sub>S</sub>SG<sub>P</sub> vs. NG-ex<sub>S</sub>NG+ex<sub>P</sub> provide evidence that cross-linguistic similarities in the overall shape of pitch accent patterns due to regional variation in L1 influence FA perception. Similar ratings between BE<sub>S</sub>SG<sub>P</sub> vs. BE<sub>S</sub>NG-ex<sub>P</sub>,

and NG-ex<sub>S</sub>SG<sub>P</sub> vs. NG-ex<sub>S</sub>NG-ex<sub>P</sub> on the other hand indicate that cross-linguistic similarities in the overall shape of pitch accent patterns due to regional variation in L1 have no effect on FA ratings. A comparison of stimuli with NG-ex<sub>P</sub> or NG+ex<sub>P</sub> validates previous findings that regional characteristics of an L2 are acquired in a naturalistic setting by L2 learners and that these characteristics reduce the degree of perceived FA.

A comparison between NG+ex<sub>S</sub>SG<sub>P</sub> vs. NG+ex<sub>S</sub>NG+ex<sub>P</sub> provides insight into the relative importance of phonetic realisation of pitch accents in judgements of FA. Higher scores for stimuli of NG+ex<sub>S</sub>SG<sub>P</sub> compared to those of NG+ex<sub>S</sub>NG+ex<sub>P</sub> suggest that the realisation of the same overall pitch pattern, i.e. the target association, is different in SG and NG+ex realisations, suggesting that the interplay between segments and prosody in the realisation of rising or falling pitch accents underlies specific regulations that are perceivable by native speakers of Belfast English. This would indicate that target alignment, that is timing of prosodic events, is important in the perception of FA. The influence of regional segmental characteristics on the perception of FA is observable in the comparison of ratings based on SG segments (SG<sub>S</sub>) vs. the remaining possible segmental levels (BE<sub>S</sub>, NG+ex<sub>S</sub>, and NG-ex<sub>S</sub>).

### 3.2 Listeners

The stimuli were presented in quasi-random order via headphones to 160 native speakers of Belfast English. The listeners were students and staff members at the University of Ulster (age range 19–53 years, average age 34) without a background in phonetics and linguistics; all with normal hearing and unpaid for their participation in the perception experiment. Due to the large number of stimuli (a total of 1760), the perception experiment was carried out with eight different sets of test stimuli. In each set 259 stimuli were tested, 215 of which were different in each session and consisted of a comparable number of stimuli per condition across the sessions. The remaining 40 stimuli were two randomly chosen stimuli per condition and judged in each of the sessions, hence by all listeners participating in the experiment. They were repeated in order to check for consistency in judgements across the individual listeners. Each session lasted about 40–50 minutes and was interrupted by a five-minute long break. Preceding the actual perception experiment, the listeners were urged to not answer immediately, but advised that they could listen to each stimulus only two times. During the instructions the listeners were informed that they would listen to acoustically modified stimuli. Each stimulus had to be judged by indicating if the sentence just heard was produced by a native or non-native

Table 4. Illustration of operational six-point scale of foreign accent (FA) rating.

Native			Non-native		
Certain	Semi-certain	Uncertain	Uncertain	Semi-certain	Certain
1	2	3	4	5	6

speaker of Belfast English. Following this forced-choice paradigm, listeners were asked for a confidence rating of their choice on a three-point scale (certain, semi-certain, and uncertain) which resulted in an operational six-point scale of FA ratings (see Table 4). This scale has been shown to be more reliable than a gradient one dimensional scale and has successfully been applied in studies on FA perception (de Leeuw, Schmid & Mennen, 2010).

Each session was preceded by a trial so that listeners could familiarise themselves with the task. The trial stimuli were repeated at different times throughout the actual experiment in order to check for consistency of judgements within individual listeners' sessions.

### Results of the perception experiment

The perception experiment was based on English utterances to allow for comparison across all four groups of speakers (BE, SG, NG+ex and NG-ex). As would be expected, there is variation within and between the speakers in the realisation of pitch accents in that not all speakers produced the same target words with the same pitch accent pattern. In addition, some of the pitch patterns were not realised at all (see the first two subsections of Section 2.4 above). Therefore, the perception experiment was based on realisations with the same pitch accent across all groups only. Thus, only target word realisations with either HL or LHL pitch patterns were considered. Only two sentences with tri-syllabic target words were realised with a rising-falling pitch accent by our speakers with the exception of the NG-ex groups. The two sentences were not produced by the same speaker. Therefore we used both speakers' utterances for the acoustic manipulation to create stimuli for the perception test. All other target realisations (LHL in bi-syllabic and HL in bi- and tri-syllabic target words) were found to be realised by all of our speakers in at least two sentences. Therefore, stimuli were created on the basis of only four sentences using the methodology detailed in Section 3.1 above. Preceding the overall analysis of FA ratings we checked for consistency of scores between the stimuli presented in all eight sessions. *T*-tests were performed for individual listeners' judgements and did not return significant results for the 40 stimuli repeated in the eight sessions nor for the trial stimuli repeated in each session. These findings allow us to assume that

judgements were consistent across the sessions and for individual listeners.

The obtained FA ratings were submitted to a mixed ANOVA with Bonferroni adjustments ( $p < .0025$ ) for multiple pair-wise comparison with *Manipulation* (with 16 levels: BE<sub>S</sub>BE<sub>P</sub>; BE<sub>S</sub>SG<sub>P</sub>; BE<sub>S</sub>NG+ex<sub>P</sub>; BE<sub>S</sub>NG-ex<sub>P</sub>; SG<sub>S</sub>BE<sub>P</sub>; SG<sub>S</sub>SG<sub>P</sub>; SG<sub>S</sub>NG+ex<sub>P</sub>; SG<sub>S</sub>NG-ex<sub>P</sub>; NG+ex<sub>S</sub>BE<sub>P</sub>; NG+ex<sub>S</sub>SG<sub>P</sub>; NG+ex<sub>S</sub>NG+ex<sub>P</sub>; NG+ex<sub>S</sub>NG-ex<sub>P</sub>; NG-ex<sub>S</sub>BE<sub>P</sub>; NG-ex<sub>S</sub>SG<sub>P</sub>; NG-ex<sub>S</sub>NG+ex<sub>P</sub>; NG-ex<sub>S</sub>NG-ex<sub>P</sub>), *Pitch accent* (with two levels: HL and LHL) and *Syllable Number* (with two levels: bi- and tri-syllabic) as within factors and *Listener* and *Gender* as between subject factors, to assess the influence of the segmental and the prosodic level of speech on listeners' scores for FA ratings. As the sphericity assumption was violated, Greenhouse-Geisser corrections were applied and are reported. The results showed a significant main effect for *Manipulation* of moderate effect size [ $F_{(8.8;2536)} = 2.46, p < .0005, \eta^2 = .462$ ]. We also found a significant interaction of moderate effect size for *Manipulation* and *Pitch Pattern* [ $F_{(6.8;988)} = 0.159, p < .0005, \eta^2 = .583$ ] and a significant interaction of a small effect size for *Manipulation* and *Listener* [ $F_{(16;961)} = 2.02, p < .0005, \eta^2 = .156$ ]. The latter is the result of variation between the 160 listeners due to different strategies employed in their decision making and would be expected. No significant interaction between *Manipulation* and *Syllable Number* was obtained so that scores for bi-syllabic and tri-syllabic target words will be combined in subsequent discussion. No significant differences between FA ratings for genders were found.

In line with previous findings (Ulbrich, 2008;), the "original" stimuli of native BE speakers (BE<sub>S</sub>BE<sub>P</sub>) received the lowest FA ratings whereas NG-ex speaker's "original" stimuli (NG-ex<sub>S</sub>NG-ex<sub>P</sub>) received the highest FA ratings (the lower FA score, the more native like is the speaker perceived, the higher the FA score, the more foreign-accented the perception of the speaker). The "original" stimuli of NG+ex (NG+ex<sub>S</sub>NG+ex<sub>P</sub>) speakers received scores in between. The FA scores for the "original" stimuli of SG speakers (SG<sub>S</sub>SG<sub>P</sub>) were only marginally lower than ratings obtained for NG-ex<sub>S</sub>NG-ex<sub>P</sub>. Figure 3 illustrates the comparison of "original" stimuli.

The interaction between *Manipulation* and *Pitch Accent* is caused by the fact that only for some of the stimuli FA scores differed significantly depending on the

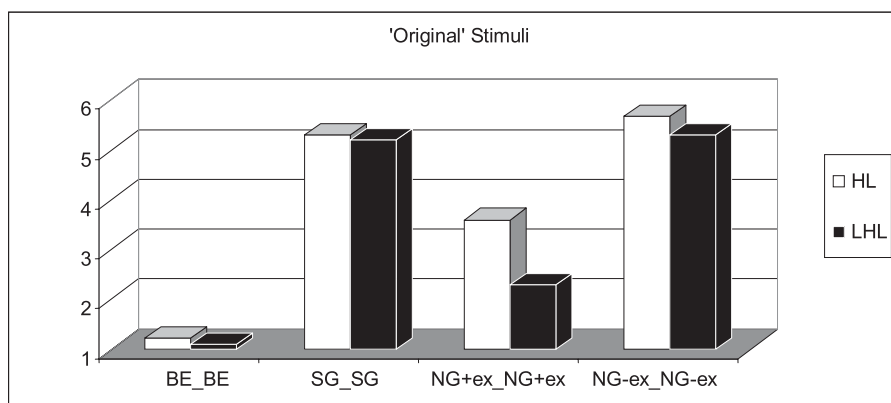


Figure 3. Foreign accent ratings for “original” stimuli of BE, SG, NG+ex and NG-ex speakers in HL and LHL realisation.

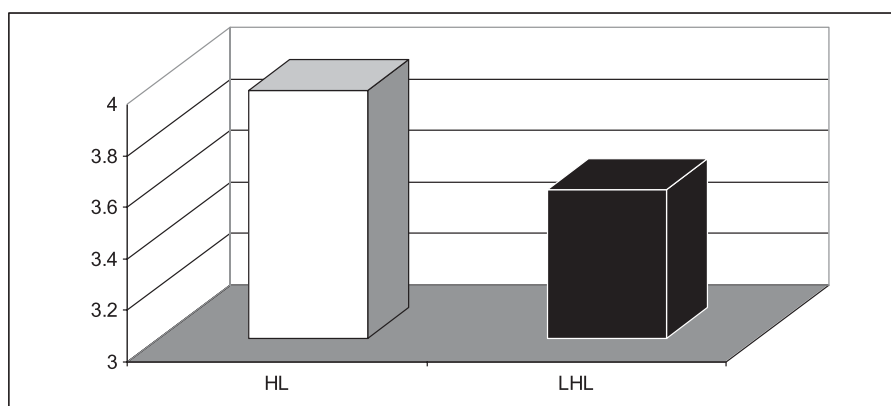


Figure 4. Foreign accent ratings pooled for HL vs. LHL stimuli across all groups of speakers (note the FA scaling 3–4 on the y-axis).

pitch pattern; i.e. falling pitch accent realisation (HL) received higher scores than rising-falling patterns (LHL), as illustrated in Figure 4.

This was found in all manipulations based on BE<sub>S</sub> and NG+ex<sub>S</sub>. Foreign accent scores for stimuli with SG<sub>S</sub> differed between HL and LHL only when combined with the prosody of BE and NG+ex speakers, i.e. SG<sub>S</sub>BE<sub>P</sub> and SG<sub>S</sub>NG+ex<sub>P</sub>; for NG-ex<sub>S</sub> only in combination with NG-ex<sub>P</sub> (see Figures 5a and 5b). Because of these differences, the results of FA ratings will be presented individually for HL and LHL pitch patterns.

Further, a comparison of FA ratings obtained for the segmental level (black columns in Figures 5a and 5b) with those obtained for the prosodic level (white columns in Figures 5a and 5b) shows that listeners are more decisive in their judgements based on the segmental level considering the range of FA scores allotted to those stimuli compared to the stimuli based on prosody.

#### **Foreign accent ratings for LHL (Figure 5a)**

Multiple pair-wise comparisons showed that NG+ex speakers received significantly lower FA scores compared to SG and NG-ex speakers not only on the basis of their

“original” stimuli as mentioned above (NG+ex<sub>S</sub>NG+ex<sub>P</sub> < SG<sub>S</sub>SG<sub>P</sub>,  $p < .0001$  and NG-ex<sub>S</sub>NG-ex<sub>P</sub>,  $p < .0001$ ) but also on the basis of segments only (NG+ex<sub>S</sub>NG+ex<sub>P</sub> < NG+ex<sub>S</sub>SG<sub>P</sub>,  $p < .0001$  and NG+ex<sub>S</sub>NG-ex<sub>P</sub>,  $p < .0001$ ). On the basis of prosody the same tendency was found (BE<sub>S</sub>NG+ex<sub>P</sub> < BE<sub>S</sub>SG<sub>P</sub>,  $p = .0062$  and BE<sub>S</sub>NG-ex<sub>P</sub>,  $p < .0001$ ; SG<sub>S</sub>NG+ex<sub>P</sub> < SG<sub>S</sub>SG<sub>P</sub>,  $p = .0052$  and SG<sub>S</sub>G<sub>P</sub>,  $p < .0001$ ; NG-ex<sub>S</sub>NG+ex<sub>P</sub> < NG-ex<sub>S</sub>SG<sub>P</sub>,  $p < .0001$  and NG-ex<sub>S</sub>NG-ex<sub>P</sub>,  $p < .0001$ ). However, the difference only approaches the significance level in some of the comparisons indicating that segments provide stronger cues than prosody for foreign accent perception. This confirms results previously presented by a number of scholars who have investigated the relative roles of prosody and segments in the perception of FA (e.g. Holm, 2008; Moyer, 1999; Munro, 1995).

By comparison, stimuli with BE segments (BE<sub>S</sub>) receive significantly lower FA ratings than those with NG+ex segments (NG+ex<sub>S</sub>) regardless of the prosodic level (BE<sub>S</sub>BE<sub>P</sub> < NG+ex<sub>S</sub>BE<sub>P</sub>,  $p < .0001$ ; BE<sub>S</sub>NG+ex<sub>P</sub> < NG+ex<sub>S</sub>NG+ex<sub>P</sub>,  $p < .0001$ ; BE<sub>S</sub>SG<sub>P</sub> < NG+ex<sub>S</sub>SG<sub>P</sub>,  $p < .0001$  and BE<sub>S</sub>NG-ex<sub>P</sub> < NG+ex<sub>S</sub>NG-ex<sub>P</sub>,  $p < .0001$ ). The comparison between stimuli based on BE

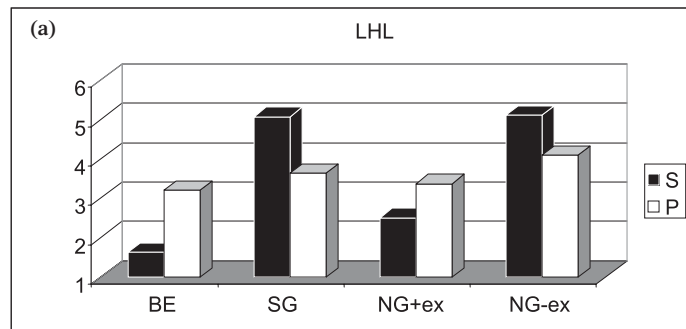


Figure 5a. Foreign accent ratings obtained for BE, SG, NG+ex and NG-ex in LHL realisation on the basis of Segments (S black) and Prosody (P white).

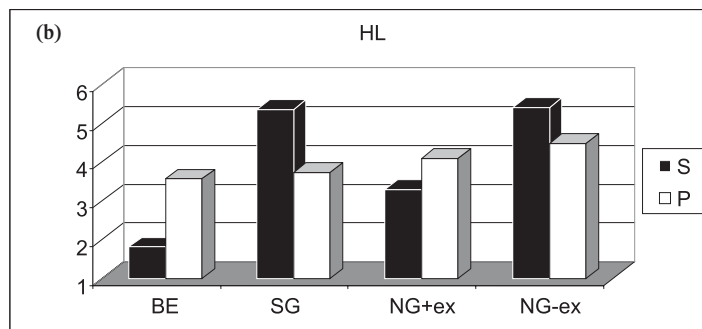


Figure 5b. Foreign accent ratings obtained for BE, SG, NG+ex and NG-ex in HL realisation on the basis of Segments (S black) and Prosody (P white).

prosody ( $BE_P$ ) with those based on NG+ex prosody ( $NG+ex_P$ ) is less clear-cut. Stimuli with  $BE_P$  attracted significantly lower FA scores than those with  $NG+ex_P$  only when they were crossed with  $BE_S$  or  $NG+ex_S$  ( $BE_S BE_P < BE_S NG+ex_P$ ,  $p < .0001$ ;  $NG+ex_S BE_P < NG+ex_S NG+ex_P$ ,  $p < .0001$ ) but not in combination with  $SG_S$  ( $SG_S BE_P$  vs  $SG_S NG+ex_P$ ,  $p = .0068$ ) and  $NG-ex_S$  ( $NG-ex_S BE_P$  vs  $NG-ex_S NG+ex_P$ ,  $p = .0031$ ).

Foreign accent ratings of stimuli with SG segments ( $SG_S$ ) and those with NG-ex segments ( $NG-ex_S$ ) did not differ significantly ( $p < .033$ ). Both stimulus groups showed a comparable pattern of ratings in that combinations with  $NG-ex_P$  received the highest ratings whereas those with  $BE_P$  and  $NG+ex_P$  received the lowest FA scores.

**Foreign accent ratings for HL (Figure 5b)**

The most striking difference between the FA ratings obtained for stimuli with rising-falling (LHL) pitch accents compared to those obtained for falling (HL) pitch patterns is that HL stimuli based on SG speaker’s prosody attracted lower scores than those based on NG+ex speaker’s prosody ( $NG+ex_P$ ) when combined with BE segments and NG+ex segments ( $BE_S SG_P < BE_S NG+ex_P$ ,  $p < .0001$ ;  $NG+ex_S SG_P < NG+ex_S NG+ex_P$ ,  $p < .0001$ ).

The same can also be seen on the basis of  $SG_S$  and  $NG-ex_S$ , in that the difference between  $SG_S SG_P$

vs.  $SG_S NG+ex_P$  ( $p = .0067$ ) and  $NG-ex_S SG_P$  vs.  $NG-ex_S NG+ex_P$  ( $p = .012$ ) is also significant. Note that on the basis of NG+ex segments,  $SG_P$  even attracted similar FA ratings compared to  $BE_P$  ( $NG+ex_S SG_P$  vs  $NG+ex_S BE_P$ ,  $p = .026$ ) (see Figure 6). When pooling stimuli according to the segmental level, FA ratings are similar to those obtained for LHL pitch accents in that BE segments ( $BE_S$ ) received the lowest, NG-ex and SG segments ( $SG_S$  and  $NG-ex_S$ ) the highest FA scores, while NG+ex segments ( $NG+ex_S$ ) are perceived as more foreign-accented than  $BE_S$  but considerably less than both  $SG_S$  and  $NG-ex_S$  ( $BE_S < NG+ex_S$ ,  $p < .0001$ ;  $BE_S < SG_S$ ,  $p < .0001$ ;  $BE_S < NG-ex_S$ ,  $p < .0001$ ;  $NG+ex_S < SG_S$ ,  $p < .0001$ ;  $NG+ex_S < NG-ex_S$ ,  $p < .0001$ ;  $SG_S = NG-ex_S$ ,  $p = .0029$ ).

A possible explanation would be that the segmental level of the two sentences produced with HL pitch patterns provided more salient segmental cues for native Belfast English speakers’ foreign accent detection, (for an overview of typical errors produced by German L2 speakers of English see Biersack, 2002). These could include among others velarisation of /l/ (which does not occur in German but does in Belfast English), vowel and syllable reductions (which follow different phonological patterns in the two language varieties), realisation of monophthongs (which are often diphthongised in Belfast English but not in Northern

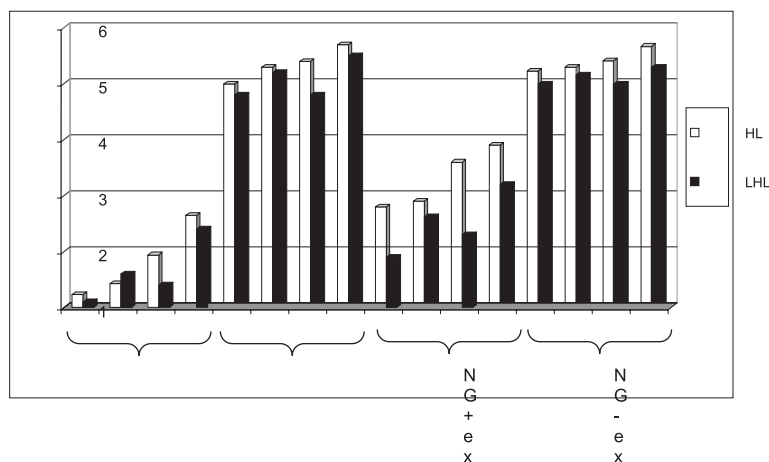


Figure 6. Foreign accent ratings for all manipulations in HL (white) and LHL (black) (four labels BE, SG, NG+ex, NG-ex in the lower panel on the x-axis refers to segmental level, the repeated four labels above refer to the prosodic level of manipulation).

Standard German), realisation of dental fricatives, labio-velar and alveolar approximants (none of which feature in NSG), as well as language-specific phonological and co-articulatory processes (i.e. directionality and dominance of assimilation). In summary, there are a number of segmental cross-language differences between Belfast English and Northern Standard German which can potentially be a source of perceptions of foreign accent. Therefore, we monotonised the four sentences produced by the 40 speakers (220 Hz and 150 Hz for female and male speakers, respectively) and asked 20 native speakers of Belfast English to rate their foreign accentedness using the same operational scale as described in Section 3.2. The FA scores were submitted to a multivariate ANOVA in order to examine differences in *FA Ratings* between the four sentences depending on *Group of Speaker* (BE, SG, NG+ex and NG-ex) and *Pitch Target* (HL, LHL). The results were not significant, showing that the segmental level for the two sentences per pitch target was perceived with a comparable degree of foreign accentedness.

In the following, key findings of the production data analysis and the results of the FA ratings obtained in the perception experiment will be summarised.

#### 4. Summary of key findings

##### 4.1 Production data

The comparison of pitch accent realisations in three groups of native German speakers across two languages revealed cross-varietal differences in both their German L1 and their English L2. The main differences between the three groups of speakers were found in the realisation of falling (HL) and rising-falling (LHL) pitch accent realisations. German speakers (NG-ex) used mostly

falling (HL) pitch accents in German as well as in English. Swiss German speakers (SG) used predominantly low-target pitch accents in German and English. However, the distribution differed depending on the language. In the German data, SG speakers realised more than 80% of the target words with low rising and rising-falling pitch accents compared to only 60% in English realisations. For L2 speakers (NG+ex) of Belfast English the reverse was the case. NG+ex speakers produced more falling (HL) pitch accents in German. In English, predominantly rising (in bi-syllabic) and rising-falling (in tri-syllabic) pitch patterns were found. Overall, NG+ex speakers of Belfast English and SG speakers realised considerably more rising and rising-falling pitch accents than NG-ex speakers in both English L2 and German L1. Another interesting similarity between the two groups of speakers is the realisation of more LH patterns in bi-syllabic words compared to LHL patterns found in the majority of tri-syllabic target words. Turning to the English data, it becomes obvious that native speakers of Belfast English (BE) practically represent the opposite to the German group (NG-ex) in that 80% of the target words were realised with rising and rising-falling pitch accent patterns.

##### 4.2 Perception experiment

As predicted, NG+ex speakers' "original" stimuli (NG+ex<sub>S</sub>NG+ex<sub>P</sub>) were perceived as more foreign-accented than "original" BE stimuli (BE<sub>S</sub>BE<sub>P</sub>), but as considerably less foreign-accented than "original" NG-ex stimuli (NG-ex<sub>S</sub>NG-ex<sub>P</sub>) by native speakers of Belfast English. No predictions were made concerning "original" SG stimuli, which received FA scores as high as "original" NG-ex stimuli. The same distribution was found in FA



ratings obtained for stimuli based on the segmental level ( $BE_S < NG+ex_S < SG_S = NG-ex_S$ ).

Overall, falling pitch accents (HL) were perceived as significantly more foreign-accented than rising-falling pitch accents (LHL). Differences between HL and LHL occurred with regard to the ranking of  $NG+ex_P$  and  $SG_P$ .  $SG_P$  yielded higher FA scores in LHL but lower scores in HL compared to  $NG+ex_P$ .  $BE_P$  stimuli attracted the lowest FA ratings and  $NG-ex_P$  the highest. However HL stimuli with  $BE_P$  and  $SG_P$  did not differ significantly and neither did LHL stimuli with  $BE_P$  and  $NG+ex_P$  (HL:  $BE_P < SG_P < NG+ex_P < NG-ex_P$ ; LHL:  $BE_P < NG+ex_P < SG_P < NG-ex_P$ ). The FA scores for all manipulations and both falling and rising-falling pitch accents are illustrated in Figure 6.

## 5. Discussion and conclusion

The analysis of the production data confirms the results of previous studies by (i) showing cross-varietal differences between Bernese German spoken in Switzerland and Northern Standard German in the realisation of nuclear pitch accents in declarative utterances (see Fitzpatrick-Cole, 1999; Ulbrich, 2004, 2005), and (ii) providing evidence for cross-linguistic differences between Belfast English and Northern Standard German (Grabe, 2002, for Belfast English). Furthermore, the study demonstrates that L2 speakers not only acquire the regionally marked pitch accent patterns of a target language (Ulbrich, 2008;), but that these L2 characteristics interfere with their L1. In this case it is the default rising-falling pitch patterns (LHL) of Belfast English which have been shown to interfere with speakers' German L1. A comparison between L2 speakers of Belfast English and native German speakers without previous exposure to Belfast English ( $NG-ex_P$ ) shows that the former produced default pitch accents of their L2 (LHL) more frequently which consequently causes a decrease of falling (HL) pitch accent realisations, the default pattern in their L1. This tendency is not only evident in the comparison between the two groups of native speakers of Northern Standard German but also in an intra-group comparison of L1 German and L2 English data produced by  $NG+ex$  speakers. The regionally marked L2 default accent (LHL) is produced in their L2 English but also in their L1 German considerably more often compared to  $NG-ex$  speakers who had not been exposed to Belfast English prior to the recordings. These findings are best accounted for by phonological usage-based theories (e.g. Johnson, 1997; Pierrehumbert, 2001) that build on a firm relationship between speech perception and production and take into consideration the impact of recent and distant linguistic experience (e.g. Bjork & Bjork, 1992; Sancier & Fowler, 1997). It seems reasonable to assume that the emergence of rising pitch accents in declarative sentences typically found in Belfast

English in realisations of L1 German speakers ( $NG+ex$ ) is the direct result of frequency of occurrence in this particular context and its associated function (Silverman, 2011).

### 5.1 LHL and HL in $NG+ex$

Applying a usage-based approach, the  $NG+ex$  speakers' data could be explained as a result of immediate exposure. Although all of the recorded  $NG+ex$  speakers living in Belfast have experienced long-term exposure to a falling pitch accent (HL) due to their growing-up in the German L1 environment, in their current language environment, i.e. English as spoken in Belfast has a rising-falling default pitch accent in declarative utterances. Therefore it seems reasonable to assume that exposure and frequency might lead to the establishment of a new categorical component in the linguistic system of  $NG+ex$  speakers.

A different though related issue arises when we try to interpret the results of the prosodic annotation in light of the FA ratings obtained in the perception experiment. Frequency effects of input from the target language for  $NG+ex$  speakers and the native language for BE speakers can account for both the production data of  $NG+ex$  speakers and the FA ratings of native Belfast English listeners. Given the nearly exclusive use of rising and rising-falling pitch accents in the variety of English spoken in Belfast, native speakers of this variety have quantitatively considerably more perceptual experience with LHL compared to HL pitch accents. This could explain why overall HL received comparably higher FA scores than LHL, indicating that listeners were more certain when rating FA in LHL stimuli compared to HL stimuli.

A possible explanation is rooted in the understanding that the distinction of falling and rising (or high and low targeted) pitch accents and their association to specific linguistic and non-linguistic functions is categorical and differs between languages and varieties. Applying Flege's (1995, p. 49) idea of "equivalence classification" referring to "a basic cognitive mechanism which permits humans to perceive constant categories in the face of the inherent sensory variability found in the many physical exemplars which may instantiate a category", we can speculate that HL corresponds to tonal categories in both languages and their varieties but that they are distinct in their phonetic implementation on the segmental level. This has previously been observed by Mennen (2004) in the study of L1 Dutch – L2 Greek interference. Hence, the perception of foreign accentedness is likely to be caused by differences in the interplay between segments and prosody (Xu, 2006), i.e. the pitch alignment (Atterer & Ladd, 2004).

This interpretation relates to Sancier and Fowler's (1997) analysis of "gestural drift" in VOT of a bilingual

speaker (e.g. Schmidt, Carello & Turvey, 1990). Adopting this idea, we assume that the L2 learners of the present study start with a HL realisation of nuclear pitch accents in their L1 German. This pattern is associated with a specific segment–prosody interplay, i.e. the synchronisation of the segmental and prosodic levels of speech. The production of pitch targets is subject to physical and mechanical properties of the larynx and its synchronisation with other articulators (e.g. Kelso, Saltzman & Tuller, 1986; Xu & Sun, 2002). Xu (2005, p. 246) argues that specific communicative functions are encoded on the basis of a limited number of articulatory primitives that are independently operable parameters (i.e. local pitch target, pitch range, strength and duration). The independence of these parameters permits a number of settings that correspond to specific communicative functions. In “syllable-synchronized sequential target approximation”, these articulatory settings are transformed into continuous movements of the acoustic signal. The differences between high target pitch accents in German varieties as well as English might be the results of language and variety specific encoding mechanisms or differences on the segmental level that define the setting of articulatory primitives. Further research, including articulatory measurements, is needed to investigate the relationship between such articulatory primitives of pitch accent realisation and varietal and/or language specific segmental characteristics.

Low targets in nuclear pitch accents of declarative sentences, by contrast, might be treated as a category that has to be newly acquired by the NG+ex learners of Belfast English and therefore its systematic encoding does not need to overcome bi-directional interference between L1 and L2. Due to extensive exposure to the low target pitch accent, LHL takes over the functional load as a default accent in declarative utterances of HL also in German L1 produced by NG+ex speakers.

That means that when exposed to the L2 linguistic environment in Belfast, two possible pitch accent realisations have to be incorporated: an L2 high target which is the default pitch target in Northern Standard German and an L2 low target without an equivalent in pitch accent realisation of L1 German. It seems reasonable to assume that HL will be associated with two possible phonetic forms (German HL and Belfast English HL) that could potentially cause interference between the L1 and L2 patterns whereas LHL will be added as a new “phonetic” category without causing interference. This would not only explain the differences in HL vs. LHL distribution between NG+ex and NG-ex speakers but it would also account for the difference between FA ratings obtained for HL and LHL. The relationship between L1 and L2 high target pitch accents is likely to cause the perception of foreign accentedness. LHL on the other hand is a phonetic category acquired on the basis of L2 input only that does not interfere with established

L1 LHL categories. These hypothetical and tentative explanations lead to further questions, for example, Are these differences dependent on the individual speaker’s language mode (see Grosjean, 2001)? How are the “new” rising or low-targeted pitch accents produced in L1 German perceived by native German speakers? Do NG-ex and NG+ex speakers differ in the phonetic implementation of falling or high-targeted pitch accents or are the interferences found in L2 data the result of reorganisation of the phonological system? What is the exact nature of synchronisation of segmental and prosodic level that causes perceived differences between NG+ex and BE? Are the differences the result of systematic L2 influence or L1 attrition? What is the extent of contribution of segmental and prosodic level and their interplay?

## 5.2 *LHL in Swiss German and Belfast English*

The influence of an L2 cannot explain the data obtained for the SG speakers. In both English and German, SG speakers produce mainly rising or low-target pitch accents, but in English they do so significantly less often. These findings are likely to be the result of classroom learning, a situation in which learners can be assumed to be exposed to a variety that is perceived as a standard variety of the acquired L2. In the present case, this variety would be Standard Southern British English (SSBE), which features HL realisations in declarative utterances; and SG learners of L2 English may have been instructed to use these default pitch accents. Previous research has produced evidence that the variety spoken by language instructors in classroom situations influences the learners’ L2 acquisition (e.g. Young-Scholten, 1985).

More intriguing are the FA ratings obtained for the SG speakers in comparison to the two other groups of native German speakers, NG-ex and NG+ex. Overall, FA ratings for SG as well as NG-ex speakers’ “original” stimuli are comparably high, hence both groups are clearly perceived as non-native speakers of Belfast English with ratings between 5 and 6 (on a six-point scale of FA ratings; see Table 4 above). However, SG speakers’ prosody only is perceived as less foreign-accented than NG-ex speakers’ prosody, regardless of the combination with the various segmental levels. This observation not only demonstrates the impact of prosody on the perception of FA but also leads to the conclusion that the question of relative contribution of prosody and segments to the perception of FA cannot be answered as straightforwardly as previously suggested. Although the findings suggest that listeners are more decisive in their judgements of stimuli based on the segmental level compared to those based on prosody, the judgements of the latter generally resemble the tendency of the former. The results seem to indicate that segments provide stronger cues compared to the prosodic level, however, prosody is nonetheless sufficiently used in FA

perception and detection as previously suggested in the literature (Holm, 2008; Moyer, 1999; Munro, 1995; Riney, Takagi & Inutska, 2005). Employing the same idea that segment–prosody synchronisation is systematic, as advocated in the interpretation of NG+ex data above, rising–falling (or low-targeted) pitch accent realisations might correspond to different encoding mechanisms in the target approximation in SG and BE. SG speakers do not establish a new “phonetic” category for the LHL pitch pattern as assumed for NG+ex speakers. From the perception point of view, the higher FA ratings for SG speakers’ LHL compared to NG+ex speakers could be due to the fact that native Belfast English listeners need to map SG speakers’ LHL – with its variety-specific articulatory properties – to the native pattern. In order to provide evidence for such an account, FA ratings would have to be obtained from native Belfast speakers that receive exposure to SG speakers’ foreign-accented speech in order to provide stronger evidence for within-category sensitivity in pitch patterns, but again we have to leave this for future research.

Furthermore, due to the very controlled experimental design, partly constrained by the requirements for stimuli creation through acoustic manipulations, it needs to be pointed out that several factors (age of first exposure, contact with L1, length of residence among others) will have to be addressed in subsequent studies. Probably most importantly, the findings have to be validated in different speaking styles including naturalistic data. What the data indisputably show is that linguistic behaviour underlies dynamic relationships with (linguistic and non-linguistic) factors that continuously influence our perception and production of language. In particular with respect to spoken languages, the phonetic detail of speech production of either a first or a second language is influenced by the speakers surrounding us (see Harrington, Palethorpe & Watson, 2000, for a scientific yet entertaining study).

## References

- Anderson-Hsieh, J., Johnson, K., & Koehler, K. (1992). The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42, 529–555.
- Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, 38, 561–613.
- Andrews, D. (1999). *Sociocultural perspectives on language change in diaspora: Soviet immigrants in the United States*. Amsterdam & Philadelphia, PA: John Benjamins.
- Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of “segmental anchoring” of f0: Evidence from German. *Journal of Phonetics*, 32, 177–197.
- Baker, W., & Smith, L. C. (2010). The impact of L2 dialect on learning French vowels: Native English speakers learning Quebecois and European French. *Canadian Modern Language Review*, 66 (5), 711–738.
- Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics in Language Teaching*, 44, 231–250.
- Bayley, R., & Langman, J. (2004). Variation in the group and the individual: Evidence from second language acquisition. *International Review of Applied Linguistics in Language Teaching*, 42, 303–319.
- Bayley, R., & Regan, V. (2004). Introduction: The acquisition of sociolinguistic competence. *Journal of Sociolinguistics*, 8 (3), 323–338.
- Benus, S., Gravano, A., & Hirschberg, J. (2007). Prosody, emotions, and . . . “whatever”. In *Interspeech*, Antwerp, 2629–2632.
- Best, C. T. (1995). A direct realist view of crosslanguage speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, pp. 171–204. Timonium, MD: York Press.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listeners’ native phonological system. *Journal of the Acoustical Society of America*, 109 (2), 775–792.
- Biersack, S. (2002). Systematische Aussprachefehler deutscher Muttersprachler im Englischen – Eine phonetisch-phonologische Bestandaufnahme [Systematic pronunciation mistakes of German natives speaking English]. In *Forschungsberichte des Instituts für Phonetik und sprachliche Kommunikation der Universität München (FIPKM)*, 39, 37–130. [M.A. thesis.]
- Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn & R. Shiffrin (eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (vol. 2), pp. 35–67. Hillsdale, NJ: Lawrence Erlbaum.
- Boersma, P., & Weenink, D. E. (2005). Praat: Doing phonetics by computer (version 5120). [Computer program.]
- Bombien, L., Mooshammer, C., Hoole, P., Kühnert, B., & Schneeberg, J. (2006). An EPG study of /k/ clusters in varying prosodic conditions in German. In H. C. Yehia, D. Demolin & R. Laboissière (eds.), *Proceedings of 7th International Seminar on Speech Production*, pp. 35–42.
- Boula de Mareuil, P., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica*, 63 (4), 247–267.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/, IV: Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310.
- Brahimi, B., Boula de Mareuil, P., & Gendrot, C. (2004). Role of segmental and suprasegmental cues in the perception of maghrebian-accented French. *Interspeech*, pp. 341–344.
- Brière, E. (1966). An investigation of phonological interferences. *Language*, 42, 768–795.
- Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, 24, 209–244.

- Byrd, D., & Choi, S. (2006). At the juncture of prosody, phonology, and phonetics – the interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In C. Fougeron, B. Kuehnert, M. Imperio & N. Vallee (eds.), *Laboratory Phonology 10*, pp. 31–60. Berlin: Mouton de Gruyter.
- Case, P., Tuller, B., Ding, M., & Kelso, J. A. S. (1995). Evaluation of a dynamical model of speech perception. *Perception & Psychophysics*, 57, 977–988.
- Chambers, J., & Trudgill, P. (1998). *Dialectology*. Cambridge: Cambridge University Press.
- Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37 (4), 466–485.
- Cho, T., & Ladefoged, P. (1999). Variations and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27, 207–229.
- Clyne, M. G. (1995). *The German language in a changing Europe*. Cambridge: Cambridge University Press.
- Crothers, J. (1978). Typology and universals of vowel systems. In J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik (eds.), *Universals of human language* (vol. 2): *Phonology*, pp. 93–152. Stanford, CA: Stanford University Press.
- Davidson, L. (2006). Phonology, phonetics, or frequency: Influence on the production of non-native sequences. *Journal of Phonetics*, 34, 104–137.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (eds.), *Handbook of second language acquisition*, pp. 313–348. Oxford: Blackwell.
- Dewaele, J.-M. (2004). Retention or omission of the *ne* in advanced French interlanguage: The variable effect of extralinguistic factors. *Journal of Sociolinguistics*, 8 (3), 333–350.
- Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40, 85–107.
- Ellis, R. (2007). Educational settings and second language learning. *Asian EFL Journal*, 9, 11–27.
- Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26, 551–585.
- Fischer-Jørgensen, E. (1968). Les occlusives françaises et danoises dun sujet bilingue. *Word*, 24, 112–153.
- Fitzpatrick-Cole, J. (1999). The alpine intonation of Bern Swiss German. *International Congress of Phonetic Sciences 1999*, pp. 941–944.
- Flege, J. E. (1981). The phonological basis of foreign accent. *TESOL Quarterly*, 15, 443–455.
- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47–65.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in crosslanguage research*, pp. 233–273. Baltimore, MD: York Press.
- Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske & A. Rohde (eds.), *An integrated view of language development: Papers in honor of Henning Wode*, pp. 217–244. Trier: Wissenschaftlicher .
- Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics*, 34, 153–175.
- Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, 185–202.
- Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers’ production and perception of English vowels. *Journal of the Acoustical Society of America*, 106, 2973–2987.
- Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, 34 (4), 409–438.
- Fowler, C., Sramko, V., Ostry, D. J., Rowland, S. A., & Halle, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics*, 36, 649–663.
- Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America*, 97, 2540–2551.
- Fujimura, O. (2000). The C/D model and prosodic control of articulatory behavior. *Phonetica*, 57, 128–138.
- Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics*, 9 (2), 198–217.
- Gass, S., & Selinker, L. (eds.) (1992). *Language transfer in language learning*. Amsterdam: John Benjamins.
- Genesee, F. (2000). Introduction: Syntactic aspects of bilingual acquisition. *Bilingualism: Language and Cognition*, 3, 167–172.
- Gilles, P. (2005). *Regionale Prosodie im Deutschen Variabilität in der Intonation von Abschluss und Weiterweisung* (Linguistik – Impulse & Tendenzen 6). Berlin & New York: de Gruyter.
- Gilles, P., & Peters, J. (eds.) (2004). *Regional variation in intonation* (Linguistische Arbeiten 492). Tübingen: Niemeyer.
- Grabe, E. (1998). *Comparative intonational phonology: English and German* (MPI Series in Psycholinguistics 7). Wageningen: Ponsen en Looien.
- Grabe, E. (2002). Variation adds to prosodic typology. In *Speech Prosody 2002*, Aix-en-Provence, pp. 127–132.
- Grabe, E. (2004). Intonational variation in urban dialects of English spoken in the British Isles. In Gilles & Peters (eds.), pp. 9–31.
- Grice, M., & Bauman, S. (2007). An introduction to intonation: Functions and models. In Trouvain & Gut (eds.), pp. 25–52.
- Gooskens, C. (1997). On the role of prosodic and verbal information in the perception of Dutch and English varieties. Ph.D. dissertation, University of Nijmegen.
- Grosjean, F. (2001). The bilingual’s language modes. In J. Nicol (ed.), *One mind, two languages: Bilinguals language processing*, pp. 1–22. Oxford: Blackwell.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. New York: Cambridge University Press.
- Gut, U. (2009). *Non-native speech*. Frankfurt/Main: Peter Lang.

- Gut, U. (2007). Foreign accent: Speaker classification I. *Lecture Notes in Computer Science*, 4343, 75–87.
- Harrington, J., Palethorpe, S., & Watson, C. I. (2000). Does the Queen speak Queen's English. *Nature*, 408, 927–928.
- Hazan, V. (2007). Second language acquisition and exemplar theory. *16th International Congress of Phonetic Sciences*, Saarbrücken, pp. 39–42.
- Hirst, D., & Di Cristo, A. (1998). *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.
- Holm, S. (2008). Intonational and durational contributions to the perception of foreign-accented Norwegian: An experimental phonetic investigation. Ph.D. dissertation, Norwegian University of Science and Technology, Trondheim.
- Ikeno, A., & Hansen, J. H. L. (2006). The effect of listener accent background on accent perception and comprehension. *EURASIP Journal on Audio, Speech, and Music Processing*, Volume 2007, Article ID 76030, 8 pages.
- Ingram, J. C. L., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, 25, 343–370.
- Isemonger, I. M. (2007). Operational Definitions of explicit and implicit knowledge: Response to R. Ellis (2005) and some recommendations for future research in this area. *Studies in Second Language Acquisition*, 29, 101–118.
- Jackendoff, R. (1997). *The architecture of the language faculty* (Linguistic Inquiry Monographs 28). Cambridge, MA: MIT Press.
- Jilka, M. (2000). The contribution of intonation to the perception of foreign accent. Ph.D. dissertation, University of Stuttgart.
- Jilka, M. (2007). Different manifestations and perceptions of foreign accent in intonation. In Trouvain & Gut (eds.), pp. 76–96.
- Johnson, K. (1997). Speech perception without speaker normalization: an exemplar model. In K. Johnson & J. Mullenix (eds.), *Talker variability in speech processing*, pp. 145–166. San Diego, CA: Academic Press.
- Jun, S.-A. (2006). *Prosodic typology*. Oxford: Oxford University Press.
- Kellerman, E., & Sharwood Smith, M. (1986). *Crosslinguistic influence on second language acquisition*. Oxford: Pergamon Press.
- Kelso, J. A. S., Saltzman, E. L., & Tuller, B. (1986). The dynamical perspective on speech production: data and theory. *Journal of Phonetics*, 14, 29–59.
- Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, 64, 459–490.
- Kim, S., & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of the Acoustical Society of America*, 125 (5), 3373–3386.
- Krakow, R. A. (1999). Physiological organization of syllables: A review. *Journal of Phonetics*, 27, 23–54.
- Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50, 93–107.
- Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, 100, 2425–2438.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.
- Ladd, D. R. (2008). *Intonational phonology* (2nd edn.). Cambridge: Cambridge University Press.
- Ladefoged, P., & Cho, T. (2001). Linking linguistic contrasts to reality: The case of VOT. *UCLA Working Papers in Phonetics*, 98, 1–9.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Malden, MA: Blackwell.
- Lakshmanan, U. (1994). Child second language acquisition of syntax. *Studies in Second Language Acquisition*, 17, 301–329.
- Leeuw de, E., Schmid, M., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant context. *Bilingualism: Language and Cognition*, 13, 33–40.
- Lepetit, D. (1987). Cross-linguistic influence in intonation. French/Japanese and French/English. *Language Learning*, 39, 397–413.
- Lowry, O. (2002). The stylistic variation of nuclear patterns in Belfast English. *Journal of the International Phonetic Association*, 32 (1), 33–42.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21, 499–546.
- Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, 26, 381–400.
- Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, 76, 190–208.
- Major, R. C. (1993). Sociolinguistic factors in loss and acquisition of phonology. In K. Hyltenstam & A. Viberg (eds.), *Progression and regression in language: Sociocultural, neuropsychological and linguistic perspectives*, pp. 463–478. Cambridge: Cambridge University Press.
- Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum.
- Major, R. C. (2004). Gender and stylistic variation in second language phonology. *Language Variation and Change*, 16, 169–188.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36, 173–190.
- McLaughlin, B. (1978). *Second-language acquisition in childhood*. Hillsdale, NJ: Lawrence Erlbaum.
- Meisel, J. M. (2001). The simultaneous acquisition of two first languages: Early differentiation and subsequent development of grammars. In J. Cenoz & F. Genesee (eds.), *Trends in bilingual acquisition*, pp. 11–42. Amsterdam: John Benjamins.
- Meisel, J. M. (2004). The bilingual child. In T. Bhatia & W. Ritchie (eds.), *The handbook of bilingualism*, pp. 91–113. Oxford: Blackwell.

- Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, 32, 543–63.
- Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, 9, 453–468.
- Moyer, A. (1999). Ultimate language attainment in L2 phonology. *Studies in Second Language Acquisition*, 21 (1), 81–108.
- Mozziconacci, S., & Hermes, D. (1998). *Study of intonation patterns in speech expressing emotion or attitude: Production and perception* (IPO Progress Report). Eindhoven: IPO.
- Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition*, 17, 17–34.
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language and Learning*, 45, 73–97.
- Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning*, 48, 159–182.
- Munro, M. J., & Derwing, T. M. (2001). Modeling perception of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, 23 (4), 451–68.
- Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
- Odling, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Ohala, J. J. (1990). There is no interface between phonetics and phonology. A personal view. *Journal of Phonetics*, 18, 153–171.
- Pavlenko, A. (2000). L2 influence on L1 in late bilingualism. *Issues in Applied Linguistics*, 11 (2), 175–205.
- Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. Ph.D. dissertation, MIT.
- Pierrehumbert, J. B. (2000). Tonal elements and their alignment. In M. Horne (ed.), *Prosody: theory and experiment: Studies presented to Gösta Bruce*, pp. 11–36. Dordrecht: Kluwer.
- Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (eds.), *Frequency and the emergence of linguistic structure*, pp. 137–158. Philadelphia, PA: John Benjamins.
- Pierrehumbert, J. B., & Beckman, M. E. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.
- Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan & M. E. Pollack (eds.), *Intentions in communication*, pp. 271–311. Cambridge, MA: MIT Press.
- Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29, 191–215.
- Riney, T. J., Takagi, N., & Inutska, K. (2005). Phonetic parameters and perceptual judgement of accent in English by American and Japanese listeners. *TESOL Quarterly*, 39 (3), 441–466.
- Saito, H. (2006). Nuclear-stress placement by Japanese learners of English: Transfer from Japanese. In Y. Kawaguchi, I. Fónagy & T. Moriguchi (eds.), *Prosody and syntax*, pp. 125–139. Amsterdam: John Benjamins.
- Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421–436.
- Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 227–247.
- Scobbie, J. (2007). Interface and overlap in phonetics and phonology. In G. Ramchand & C. Reiss (eds.), *The Oxford handbook of linguistic interfaces*, pp. 17–52. Oxford: Oxford University Press.
- Selkirk, E. O. (1995). Sentence prosody: Intonation, stress and phrasing. In J. A. Goldsmith (ed.), *The Handbook of phonological theory*, pp. 550–569. Oxford: Blackwell.
- Silverman, D. (2011). Usage-based phonology. In B. Botma, N. C. Kula & K. Nasukawa (eds.), *Continuum companion to phonology*, pp. 369–395. London & New York: Continuum.
- Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25, 1–24.
- Thomas, M. (2005). Theories of second language acquisition: Three sides, three angles, three points. *Second Language Research*, 21 (4), 393–414.
- Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1–30.
- Trofimovich, P., & Baker, W. (2007). Learning prosody and fluency characteristics of L2 speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics*, 28, 251–276.
- Trofimovich, P., Gatbonton, E., & Segalowitz, N. (2007). A dynamic look at L2 phonological learning: Seeking psycholinguistic explanations for implicational phenomena. *Studies in Second Language Acquisition*, 29, 407–448.
- Trouvain, J., & Gut, U. (eds.) (2007). *Non-native prosody: Phonetic description and teaching practice* (Trends in Linguistics: Studies & Monographs). Berlin: Mouton de Gruyter.
- Tsukada, K., Birdsong, D., Mack, M., Sung, H., Bialystok, E., & Flege, J. E. (2004). Release bursts in English word-final voiceless stops produced by native English and Korean adults and children. *Phonetica*, 61, 67–83.
- Ulbrich, C. (2004). A comparative study of declarative utterances in Swiss and German standards. In Gilles & Peters (eds.), pp. 99–122.
- Ulbrich, C. (2005). *Phonetische Untersuchungen zur Prosodie der Standardvarietäten des Deutschen in der Bundesrepublik Deutschland, in der Schweiz und in Österreich* (Hallesche Schriften zur Sprechwissenschaft und Phonetik 16). Frankfurt/Main: Peter Lang.
- Ulbrich, C. (2008). Acquisition of regional pitch patterns in L2. *Speech Prosody 2008*, 575–578.
- Vaissière, J., & Boula de Mareuil, P. (2004). Identifying a language or an accent: From segments to prosody.

- Proceedings MIDL (Modelling for the Identification of Languages) Workshop*, Paris, pp. 1–6.
- van Bezooijen, R., & Gooskens, C. (1999). Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology*, 18 (1), 31–48.
- van Leyden, K., & van Heuven, V. J. B. (2003). Prosody versus segments in the identification of Orkney and Shetland dialects. *15th International Congress of Phonetic Sciences*, Barcelona, pp. 1197–1200.
- Walley, A. C., & Flege, J. E. (1999). Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*, 27, 307–332.
- Wichmann, A. (2000a). *Intonation in text and discourse*. London: Longman.
- Wichmann, A. (2000b). The attitudinal effects of prosody, and how they relate to emotion. In R. Cowie, E. Douglas-Cowie & M. Schröder (eds.), *Proceedings of the ISCA Workshop on Speech and Emotion*, Belfast, Ireland, pp. 143–148.
- Willems, N. (1982). *English intonation from a Dutch point of view*. Dordrecht: Foris.
- Williams, L. (1979). The modification of speech perception and production in second language learning. *Perception & Psychophysics*, 26, 95–104.
- Williams, L. 1980. Phonetic variation as a function of second-language learning. In G. Yeni-Komshian, J. Kavanagh & C. Ferguson (eds.), *Child phonology* (vol. 2): *Perception*, pp. 185–215. New York: Academic Press.
- Wode, H. (1978). The beginnings of L2-phonological acquisition. *International Review of Applied Linguistics in Language Teaching*, 16, 109–124.
- Wode, H. (1992). Categorical perception and segmental coding in the ontogeny of sound systems: A universal approach. In C. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological development: Models, research, implications*, pp. 605–631. Baltimore, MD: York Press
- Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication*, 46, 220–251.
- Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 111: 1399–1413.
- Young-Scholten, M. (1985). Interference reconsidered: The role of similarity in second language acquisition. *Selecta*, 6, 6–12.
- Zubizarreta, M. L., & Vergnaud, J.-R. (2005). Phrasal stress and syntax. In M. Everaert & H. van Riemsdijk (eds.), *The Blackwell companion to syntax* (vol. III), pp. 522–568. Oxford: Blackwell.