44. Sentence prosody in a second language

Jürgen Trouvain
Bettina Braun

44.1 Introduction

This chapter aims to give an overview of the state of the art on the research of sentence prosody (i.e. the prosodic properties beyond the word level) in a second language, L2. This is complemented by two aspects of perception: L2 speech or foreign accents by listeners of the native language (L1) and L1 speech by L2 listeners (for L2 word prosody see chapter 39). By 'L2' we refer to any second or non-native language that is acquired after childhood, and by 'L1' the first or native language(s). We use the term 'native speaker' to refer to speakers of the learners' target languages. L2 *sentence* prosody is still a relatively underexplored field of L 2 acquisition, as evidenced by a lack of discussion on this topic in various handbooks on L2 acquisition (e.g., Doughty and Long 2005; Gass and Mackey 2012; Herschensohn and Young-Scholten 2013; Ritchie and Bhatia 2009). In recent years research on L2 sentence prosody has been boosted by the collection, annotation and provision of phonetic learner corpora (Trouvain, Zimmerer, Möbius, Gósy and Bonneau 2017). Examples with scripted and unscripted speech include the LeaP corpus (Gut 2012), the COREIL corpus (Delais-Roussarie and Yoo 2011), or the AixOx corpus (Herment, Loukina and Tortel 2012). However, with annotation of learner data we face a problem that is inherent in the annotation of L2 research and in prosodic annotation in general, the choice of appropriate measures or categories: annotation can be done at the purely acoustic level, with category labels of the L1, with category labels of the L2, with new interlanguage categories, or with 'error' categories, as discussed by Ramírez Verdugo (2005), Bartkova, Delais-Roussarie and Santiago Vargas (2012) and Albin (2015). The choice of coding has consequences for reliability, validity and the comparability of results across studies. Existent research is typically concerned with topics on intonation (e.g. types of pitch accents, their location and function, types of boundary tones and their functions, prosodic phrasing) or on timing (e.g. rhythm, tempo, pauses, fluency); research combining L2 melodic and timing aspects is still sparse, as noted by Mennen and de Leeuw (2014).

In what follows, we will review research on intonation in section 40.2 and research on timing in section 40.3, discuss work on perception of L2 sentence prosody in section 40.4,

and conclude with a discussion on challenges facing in this field. As in other areas of L2 acquisition, parts of L2 prosody become more native-like with earlier age-of-acquisition (see e.g., Huang and Jun (2011 for different prosodic features across groups), higher proficiency and fluency in L2 (e.g., Gut 2009b; Swerts and Zerbian 2010), and for many though not all individuals over time (e.g., Wieden 1993). Although we cannot discuss all these extra-linguistic aspects in detail, we will mention them where appropriate in our review.

44.2 Intonational aspects of L2 sentence prosody

In this section we review two linguistic functions of prosody, the prosodic marking of information structure (40.2.1) and questions (40.2.2) and then turn to the acquisition of phrasing (40.2.3) as well as the phonetic implementation of prosodic events (40.2.4). We end this section with an overview of the prosodic marking of non-linguistic attributes (40.2.5).

44.2.1 Prosodic marking of information structure

Information structure can have several partitions: focus-background, topic-comment or given-new (Krifka 2008, see also chapter 30). Focus signals the presence of alternatives, topic links an utterance to the prior discourse, while the given-new partition refers to information that either has been mentioned before or not. An additional aspect to the various partitions is contrastive information, i.e. information that contrasts with or corrects prior information or assumptions.

Information structure is signalled via various linguistic devices, i.e., lexical markers, syntactic operations, and prosody (e.g., Steindel Burdin et al. 2014; Vallduví and Engdahl 1996). Languages do not only differ in the relative importance of these devices but also in the use of prosody. For example, languages can differ in the exact use of prosodic prominence for focus marking even if they are similar in marking of phrase-level prosodic prominence. Languages can also differ whether attenuation of prosodic prominence is used to mark givenness and post-focal information.

Languages can be grouped according to how they mark phrase-level prosodic prominence: language can mark the *head of the phrase* (e.g., English, using different pitch accents), the *edge of the phrase* (e.g., Korean, using tonal marking at prosodic edges), or *both* (e.g,. French, Japanese, cf. Sun-Ah Jun 2014). In head-prominence languages (e.g., English,

German, Dutch) a pitch accent is assigned to the word in focus (see (1) for a pitch accent on *everybody*). In broad focus (i.e., a response to a question like "What's new?"), the pitch accent typically falls onto a so-called focus exponent that is defined syntactically (*prosody* in the example sentence in (1), cf. Ladd (2008). Acoustically, the result is a salient f0-movement on the focus exponent (i.e. word/s in focus or focus constituent), higher intensity, and longer duration, but also a higher probability of pauses following the focused constituent (e.g., Arnhold 2016). F0 stands for the frequency of vocal fold vibration, whose perceptual correlate is pitch. In some lexical tone languages (e.g., Mandarin Chinese), which are also head-prominence marking languages, the focused constituent is realised by an increase in f0-range of the lexical tone, followed by compression of the f0-range in the post-focus constituent (e.g., Xu 1999), but other tone languages show pre-focus pitch raising and pitch compression from the focused word onwards (e.g., Bemba, cf. Kula and Hamann 2017). Other tone language show no prosodic focus marking at all (Downing and Rialland 2017, p. 7). In contrast, in other head-prominence languages like Italian, Catalan and Spanish, information structure is mainly marked via syntactic operations, such as dislocations or syntactic movement (e.g., Büring 2009; Dehé, Feldhausen and Ishihara 2011; Sun-Ah Jun 2005; Vallduví and Engdahl 1996). These languages often do not deaccent given information, like *prosody* in the answer of example (1), and produce an additional accent on *prosody*. In edge-prominence languages (e.g., Korean), on the other hand, a focus constituent is marked by the initiation of a new prosodic phrase, into which the post-focus constituent is integrated.

(1) Question: Who is interested in prosody?

Answer: Everybody is interested in prosody.
(Stressed syllables are underlined, pitch accent is displayed in the schematic f0 contour)

Learners of the head-prominence languages outlined above are faced with the questions of where to locate pitch accents and where to implement phrase breaks (let alone syntactic operations, which are not dealt with here, cf. Hertel 2003; Zubizarreta and Nava 2011). In contrast, learners of edge-prominence languages face the challenge of producing phrasing accordingly. Learners of both types of languages need to acquire the phonetic realisation of

accents and phrase breaks (see section 40.2.4).

The acquisition of the prosodic realisation of focus in head-prominence languages has been studied in great detail, mostly with English as the target language. Learners whose L1 is a head-prominence language but marks focus syntactically or is a tone language appear to encounter more difficulty than learners whose L1 is similar to English in prosodic system and in the use of accent placement for focus marking. For example, Ramírez Verdugo (2006) reported that Spanish learners of English overgeneralized broad focus realizations to contrastive focus contexts, i.e., they did not produce a rising-falling accent on *everybody*, but an accent on *prosody* in (1). In utterances in which the accent location was correct, learners often differed in accent type from native speakers: they produced more rising accents, while native speakers of English produced more falling accents. In another study, Spanish learners of English were reported to insert pauses after the focused constituent, a strategy that was not present in native English speakers (Ortega-Llebaria and Colantoni 2014). Similarly, learners from tone languages have difficulties in accent location (Baker 2010 for Mandarin speakers of English; Swerts and Zerbian 2010 for Zulu speakers of English). German learners of English showed the same accent placement as native speakers of English in a variety of focus conditions but differed in (a) accent type from native speakers of English in using rising accents more frequently and (b) in phonetic implementation (O'Brien and Gut 2010). Learners from edge-prominence languages (Korean, Japanese) were more accurate than Mandarin learners of English in placing the pitch accent in different focus conditions (Baker 2010), but there are studies showing wrong accent placemnt in Japanese learners of English, cf. Saito (2006)..

There is, however, limited research on the acquisition of focus in edge- and head/edge-prominence languages. Current findings suggest that learners from an edge-prominence language can learn the target pattern if the target languages accent accents on structural principles, instead of information structural reasons. For example, in French, accent distribution is largely independent of whether information is new, given, or contrastive (Krahmer and Swerts (2001), advanced Dutch learners of French were found to largely produce the French pattern (Rasier and Hiligsmann 2007). For the head/edge-prominence language Japanese, Swedish learners of Japanese (i.e., they put too much emphasis (f0-scaling) on the topic constituent and too little emphasis on the focus; Nagano-Madsen 2014).

In target languages that mark information structure syntactically, the main challenge lies in the acquisition of the correct word order (e.g., Hertel 2003) and – for learners from

certain head-prominence languages – the suppression of prosodic highlighting. Turco, Dimroth and Braun (2015), showed that Dutch and German learners of Italian overly marked affirmative polarity contrast in their L2 productions, either by lexical markers (Dutch) or by a rising-falling pitch accent (German), markings that are not present in Italian L1. Likewise, English learners of Spanish were shown to use higher intensity and pitch to mark focus and contrast while native speakers of Spanish used syntactic and lexical markers (Kelm 1987).

Finally, we turn to deaccentuation of given information. English, German and Dutch are typical deaccenting languages, while Italian and Spanish are not (Brown 1983; Ladd 2008). Learners from languages without deaccentuation of given information often fail to deaccent given referents in an L2 deaccenting language (cf. Gut and Pillai 2014 for Malaysian Malay speakers of English; Nguyễn, Ingram and Pensalfini 2008 for Vietnamese learners of English; Swerts and Zerbian 2010 for Zulu speakers of English; Ueyama and Jun 1998 for Korean and Japanese learners of English). In contrast, learners with native languages that deaccent given information overuse deaccentuation in an L2 that does not deaccent (e.g., Rasier and Hiligsmann (2007) on Dutch learners of French). A related aspect to the issue of deaccenting given information is post-focus compression (PFC), a mechanism whereby constituents following the focused one exhibit shorter durations, a more compressed f0-range and lower intensity (Eady, Cooper, Klouda, Mueller and Lotts 1986; Hindi: Patil U 2008; Finnish: Vainio and Järvikivi 2007; Mandarin: Xu 1999). Recent production data show that Taiwanese learners of Mandarin Chinese, i.e., speakers whose L1 does not have PFC (Southern Min, cf. Xu, Chen and Wang 2012) but the target language does, do not consistently produce PFC in their L2 Mandarin. It is rather the case that the correct acquisition of PFC seems to be guided by L2 use (Y. Chen, Y. Xu, S. Guion-Anderson 2014).

In sum, much research has concentrated on the prosodic marking of focus, given vs. new, and contrastive information. Existent studies suggest a strong influence of L1 and provide some evidence for successful learning in certain L1-L2 pairings. However, there is comparatively little work on the acquisition of topic marking in L2. Also, there is a need for research of a wider range of L1-L2 pairings to better study the underlying mechanisms.

44.2.2 Prosodic marking of questions

Question forms include polar (yes/no) questions, constituent (*wh*)-questions, alternative and tag questions. Depending on the language, neutral polar questions are marked syntactically (e.g. German, English), by particles (e.g., Urdu, Japanese), or purely prosodically (e.g.,

Italian, Basque). In many languages, one can use a declarative syntax to ask a question (2a), even though the pragmatic effect may differ from a syntactically marked question (2b). Constituent questions are lexically marked by a question word, alternative questions by alternatives and tag questions by tags (Bartels 1999). In many languages, the intonation of questions is quite variable (Braun, Dehé, Neitsch, Wochner and Zahner 2018; Hedberg, Sosa and Fadden 2004; Hedberg, Sosa, Gürgülü and Mameni 2010; Kohler 2004), which makes it hard to establish what the native language (intonational) grammar is.

(2) a. You study prosody?

b. Do you study prosody?

Past work on L2 prosodic marking of questions shows that learners differ from native speakers in (a) prosodic realization or (b) in the distribution of realizations. With respect to English polar questions, which often end in a high-rise (Quirk, Greenbaum, Leech and Svartvik 1985), Greek learners of English transferred the typical polar question contour of their L1, which is a rise-fall (Arvaniti, Ladd and Mennen 2006), to English and also placed the nuclear accent on the verb instead of the argument of the verb as native speakers of English do (Kainada and Lengeris 2015). The use of a falling contour in English polar question was also reported for Thai and Spanish learners of English (Wennerstrom (1994). McGory (1997) compared declaratives and polar questions spoken by beginning and advanced Mandarin and Korean learners of English to those produced by native speakers of English. Beginners produced the final high boundary tone which is often used in English polar questions, but failed to produce the nuclear accent with a low tone and instead produced a high-falling accent, which typically occurs in English declarative sentences. In English, question tags may be falling or rising (Dehé and Braun 2013), the pattern being influenced by the polarity of the tag (positive or negative) and the position of the tag in the speaker's turn. Spanish learners of English were found to use a rising tag irrespective of polarity and position of the tag (Ramírez Verdugo and Romero Trillo (2005). Mexican Spanish learners of French generally had a higher proportion of high-rising boundary tones than French native speakers, irrespective of sentence type (Santiago Vargas and Delais-Roussarie (2012).

To sum up, many factors influence the intonational realisation of questions in different languages, such that the learner has to make sense of a pattern that is obscured by a

lot of variability. This may make it hard to approach the distributions of target intonational patterns and to figure out the factors that affect the intonation pattern in the non-native language.

44.2.3 Prosodic phrasing

Depending on the language, the prosodic hierarchy distinguishes between prosodic phrases at different levels: accentual phrases (e.g., Japanese, Korean, French), intermediate phrases (e.g., English, German, often termed minor phrases) and intonational phrases (major phrases), (e.g., Nespor and Vogel 2007). Learners have been shown to differ from native speakers in the number of phrase breaks they make and the tonal marking of the phrasal breaks. For example, in non-final position, e.g., after the subject-NP such as *protection* in (3), French learners of English produced more phrase breaks than native speakers of English and ended phrases mostly with rising contours (Herment, Ballier, Delais-Roussarie and Tortel 2014), while native speakers of English marked these minor phrases by f0-falls (Horgues 2013). On the other hand, learners of French from languages without accentual phrases (e.g., Mexican Spanish) produced fewer phrases than native speakers of French (Santiago Vargas and Delais-Roussarie 2012).

(3) The idea of a good protection / is to guarantee that your computer doesn't get infected by a virus/ (slashes indicate phrase breaks, example from Horgues 2013)

In Korean, phrasing may distinguish between a polar question and a constituent question reading. S.-A. Jun and Oh (2000) tested the acquisition of phrasing in minimal pairs, recording Korean natives and American learners of Korean with varying proficiency. Phrasal grouping improved with increased proficiency, but learners generally produced more stress accents, a type of accent which is absent in Korean. In sum, these studies suggest that phrasal marking is likely transferred from the native language, with proficiency being a modulating factor.

44.2.4 Phonetic implementation of pitch accents and boundary tones

Previous research has shown that L2 prosody differs from that of the target language in the alignment of f0-peaks and f0-troughs of rises, peak scaling and global f0-range. With respect

to alignment, German learners of English were shown to align accentual tones later than native speakers of English, due to influence of L1 dialect or regional accent (Atterer and Ladd 2004; Gut 2009b; Ulbrich 2013). A later alignment of high accentual tones is reported for Japanese and Spanish beginning and advanced learners of American English (Northern Virginia/Washington) by Calbert Graham (2018). However, learners are not always later in their alignment: Dutch learners of Greek, for instance, produced rising accents with earlier alignment than Greek native speakers (Mennen 2004), a likely transfer from their L1.

In terms of scaling of pitch accents, Mandarin L2 speakers of American English produced accented words with higher f0-peaks than native speakers (Y. Chen, Robb, Gilbert and Lerman 2001). Thai, Japanese and Spanish learners of English increased f0 as much as the English native speakers to mark focal information but they showed less reduction of f0 on non-focused information than the native speakers (Wennerstrom (1994).

Regarding the production of the nuclear tune (last pitch accent of the phrase plus the following boundary tone), native speakers of German truncate falls if there is little sonorant material (i.e., they stop the f0-movement earlier in names like *Shift* as compared to names like *Sheafer*), while speakers of English compress them (i.e. they realise the full f0-movement in less time) (Grabe 1998). German learners of English were found to transfer the truncation of falling accents when there is limited sonorant material, whereas English learners of German could correctly truncate the falling nuclear contours in L2 German ((Zahner and Yu 2019).

Finally, mixed findings have been reported for the production of f0-range in L2. Some studies observed a narrower f0-range for learners compared to native speakers (e.g., Kainada and Lengeris 2015 for Greek learners of English), whereas other studies found the reverse (Aoyama and Guion 2007 for Japanese learners of American English; Santiago Vargas and Delais-Roussarie 2012 for Mexican Spanish learners of French). A third group of studies finds no differences (Wennerstrom 1994 for Thai, Japanese and Spanish learners of English; Zimmerer, Andreeva, Jügler and Möbius 2015 for French learners of German, and German learners of French). Possibly, there are a number of other critical factors at play, such as speaker idiosyncrasies, the specific L1/L2 pairing, non-linguistic factors (e.g., level of uncertainty) and pragmatic context, as well as the speech task.

44.2.5 Prosodic marking of non-linguistic aspects

Besides linguistic meaning, prosody is often used to signal the speaker's emotion, epistemic

belief and certainty in questions (e.g., Domaneschi, Romero and Braun (2017), commitment (Truckenbrodt 2012) or attitude towards the proposition (Crystal 1969; Ladd 1980; O'Connor and Arnold 1973; Pierrehumbert and Hirschberg 1990; Wells 2006). Research on the use of prosody in marking non-linguistic meaning in L2 is rare. Existent work suggests that this aspect of L2 prosody also poses challenges to L2 learners, especially in the absence of explicit instruction, possibly due to differences between L1 and L2. For example, Ramírez Verdugo (2005) investigated how Spanish learners of English and native speakers of English realised uncertainty. While native speakers of English used a fall-rise to mark uncertainty and falling contours to mark certainty, the learners mostly used falling contours with a narrow f0-range in both the certainty and uncertainty condition, which made it hard to perceive the contrast between the two conditions. Furthermore, A. Chen and De Jong (2015) examined the prosodic realisation of sarcasm in advanced Dutch learners of English. They found that learners did not sound sarcastic to native speakers of English but sounded somewhat more sarcastic to native speakers of Dutch. However, learners could produce more sarcastic-sounding prosody after brief training (Smorenburg, Rodd and Chen 2015).

44.3 Timing phenomena in L2 sentence prosody

This section deals with the timing phenomena, i.e., rhythm, tempo and fluency, in L2 speech, which can have a dramatic impact on the perceived degree of foreign accent as well as the comprehensibility of L2 speech.

44.3.1 Rhythm

The definition of rhythm is quite problematic (see also chapter 10). There is a classic perception-based division of languages into stress-timed, syllable-timed and mora-timed languages, depending on which of these entities defines approximately isochronous intervals (Ladefoged 1975; Pike 1945). Some studies on L2 speech used metrics for the quantification of speech rhythm, such as the ‘pairwise variability indices’ and interval measures of vowels and consonants, together with rate-normalised interval measures (e.g., Li and Post 2014; Ordin and Polyanskaya 2015; White and Mattys 2007). In these studies, some of these rhythm measures reflected differences between L1 and L2 speech for read sentences, albeit not in a

consistent way across the studies.

However, in view of the repeated lack of evidence for the classic perception-based rhythmic categories, it is important to integrate (i) tempo and prosodic phrasing, (ii) prominence structure, and (iii) segmental reductions in research on L2 rhythm, because these properties typically work in tandem. For example, the faster we speak, the fewer phrases and pitch accents we tend to produce (e.g., Trouvain and Grice 1999), along with more segmental and syllable reduction as often observed in spontaneous speech (Barry and Andreeva 2001; Engstrand and Krull 2001). Barry (2007) and Gut (2009a) examined temporal and prominence structure instead of adopting a rhythm measure in L2 German and English. They found that in L2 speech, especially by learners at lower proficiency levels, the durational relation between strong and weak syllables typically fell short of the L1 norm, whereas more advanced learners were able to produce patterns more similar to the target patterns.

There are still other sources that are responsible for a deviant rhythm, some of which reside at the word level (wrong word stress, inappropriate usage of vowel reduction and deletion, insertion of epenthetic vowels) or at the utterance level (wrongly placed sentence accent and prosodic phrase breaks, missing or inappropriate linking between words). These aspects are important when explaining L2 speech rhythm and when studying how L2 speakers acquire and master appropriate rhythmical patterns.

44.3.2 Tempo and pauses

L2 speech is often characterised by a slower tempo than L1 speech, due to a slower articulation rate and more and longer pauses (Pürschel (1975), Wiese (1983) for German learners of English, Trofimovich and Baker (2006) for Korean learners of English). When considering stretches of speech beyond a single utterance, it is useful to distinguish between the rather general terms such as ‘speaking rate’, ‘speech rate’ and ‘tempo’, which typically include pauses, and the term ‘articulation rate’ which excludes pauses (Trouvain 2004). These measures are expressed in linguistic units per time unit, e.g., syllables per second (syll/s), words per minute (wpm), mean segmental duration, or phones per second. The most widespread metric for speech tempo seems to be syll/s (see Trouvain 2004). Note, however, that syll/s can be problematic in cross-linguistic studies. For example, languages like English and German tend to omit entire syllables; it remains unclear whether syllables are to be considered at the underlying (phonological) level or at the phonetic surface in these languages. Moreover, differences in syllable complexity may cause biases, such that a German speaker’s speaking

rate in syll/s in German with rather complex and hence longer syllable durations may be slower than her/his rate in an L2 with a less complex syllable structure (e.g., Trouvain and Möbius 2014). Words per minute have the advantage of easy counting, but have the disadvantage of cross-linguistic differences in word length and coarse granularity, while segments per second have a more fine-grained granularity, but are prone to omissions, harder to define, and are more time-consuming to count.

Various studies show that, with increasing proficiency, the tempo of L2 speaker becomes faster, and thus more similar to the tempo of the learner's L1 and target language (Trouvain and Möbius 2014). On the perceptual side, it seems common that beginning L2 learners have the impression that L1 speakers speak at an extremely fast tempo (Abercrombie 1967, p. 96). Schwab and Grosjean (2004) investigated the relative tempo perception of L1 and L2 in Swiss-German learners of French and found that the measured speech rate in L2 positively correlated with the perceived speech rate in L2 by L2 learners and negatively correlated with speech comprehension in L2.

As mentioned above, pause is a critical concept to the discussion of tempo and related terms. Pauses can be defined as phases of the speaking process in which the articulatory and phonatory activity is interrupted. However, there is no generally accepted threshold when a silence should count as a pause (distinguishing it from the closure phase of plosives, for example). Pauses are sometimes divided into 'silent' and 'filled' pauses. The latter correspond to filler syllables like 'erm' and 'uh'. The so-called 'silent' pause often contains inhalation noise, aka breath pauses, which are usually longer than silent pauses (Grosjean and Collins 1979; Trouvain, Fauth and Möbius 2016). However, silences or breath noises are not required for the perception of a perceived pause due to the syntactic expectation together with cues such as final lengthening and the shape of the nuclear contour (Butcher 1981). L2 speech tend to have more pauses and longer pauses. But a study on Dutch L1 and L2 speech showed that these L2 pausing characteristics mainly concerned pauses within utterances but not between utterances (de Jong (2016).

44.3.3 Fluency

Fluency is an important and often-mentioned concept in the assessment of L2 proficiency (Council of Europe, 2001). Higher fluency in an L2 are associated with higher proficiency levels. Although we may have an intuitive idea of what fluency is, it is not easy to define it. Production fluency is based on a speech signal, which can be used for quantitative

measurements. However, there is no agreement on the best parameter for production fluency (e.g., de Jong 2016; Gut 2009a; Raupach 1980). Measures that are frequently mentioned include the two tempo metrics articulation rate and speaking rate, but also mean length of run (run = inter-pause stretch), the 'phonation/time ratio' (ratio of articulation time to total speaking time), the number and the duration of unfilled pauses, and the number of filled pauses and other disfluencies (de Jong 2016; for details see Gut 2009a, p. 89ff). In contrast to production fluency, perceptual fluency is based on the fluency assessment of listeners (mostly native speakers of the target language). Quite often the measurements of production and perceptual fluency differ.

Fluency cannot be considered without disfluencies. In spontaneous speech, there is a number of markers of production disfluency. Formulations can be discontinued and the re-start can lead to a repair. The repair phase consists of the syllables to be repaired (reparandum), followed by the interregnum (or editing phase) after the interruption point, and terminated with the reparans as the actually repaired sequence (Levelt 1983). The interregnum can contain an explicit editing term, e.g. 'no, I mean', or in many cases silent pauses and/or filled pauses. Interestingly, filled pauses hardly occur in read speech (Duez 1982; Trouvain et al. 2016), but they are very common in spontaneous speech (e.g., Cucchiarini, Strik and Boves 2002; Duez 1982). Filled pauses can also occur in articulatory phases without any interruption and without any silence. In those cases they may not be considered a disfluency, but a means of fluency, sometimes also called 'fluenceme' (Götz 2013). In contrast to the negative associations that disfluencies trigger regarding the flow of spoken information, fluenceme may help the listener in speech comprehension. For example, they may elicit prediction of less accessible referents and may shift the attention of listeners to the upcoming information (e.g., Corley, MacGregor and Donaldson 2007). Although disfluencies can be observed in both L1 and L2 speech, there is evidence that their beneficial effect for the speech comprehension process is present for L1 disfluent speech but not for its L2 counterpart (e.g., Bosker, Quené, Sanders and de Jong 2014).

44.4 Perception of L2 sentence prosody

44.4.1 Perception and interpretation

In this sub-section we review how L2 speakers perceive differences in the intonational form of the target language (e.g., discrimination of contours, determination of accent location) and

how they interpret the role of prosody in signalling information structure, questionhood, disambiguating syntactic ambiguities and paralinguisic meaning.[1]

Baker (2010) tested the perception of accent location and the interpretation of information structure of Korean learners of English and found that the learners were as good as the native controls in determining accent location but had poorer performance than the English controls in the interpretation of information structure, suggesting influence from L1.

There are very few studies on the perception of questions vs. statements in the L2 in non-tonal L2s. Puga, Fuchs, Setter and Mok (2017) tested the ability of German learners of English to match the intended intonation pattern to a number of sentence types and functions (polar questions, tag questions, statements, sarcasm). They found that the German learners did not differ from the English controls for polar questions and statements, but were less accurate than the native controls for tag questions and sarcasm, possibly due to L1 transfer. For instance, compared to German, tag questions are more widely used in English and display greater variability in syntactic form and prosody (Dehé and Braun 2013), which apparently has consequences for L2 acquisition. Yang and Chan (2010) tested the perception of question vs. statement interpretation in English learners of Mandarin Chinese. They reported that the learners at all proficiency levels made most errors when statements ended with a syllable that had Tone 2 (f0-rise) or when questions ended in Tone 4 (high-falling), i.e., when the final f0-contour mismatched the typical contours of polar questions (rising) and statements (falling). These results also provide evidence for L1 transfer. Liang and Van Heuven (2007) tested the perception of question vs. statement in Mandarin Chinese by three learner groups, two from dialects of Chinese (Nantong and Changsha dialect) and one from a non-tone language (Uygur, an Altaic language). For the intonation task, participants had to indicate whether the utterance was a statement or question (in utterance in which tones was constantly a high-level tone). The learners from the non-tone language were more sensitive to the statement-question contrast than the learners from a tone language (who, in turn, were more correct in a separate tone recognition task).

Regarding the use of prosody to disambiguate sentence meanings in L2, Cruz-Ferreira (1989) examined English learners of Portuguese and Portuguese learners of English in their L1 and L2. She presented the participants with sentences whose meaning can be

[1] Relevant psycholinguistic studies (e.g., Akker and Cutler 2003; Braun and Tagliapietra 2011; Lee and Fraundorf 2016) and neurolinguistic studies (e.g., Nickels, Opitz and Steinhauer 2013; Nickels and Steinhauer 2018) are not reviewed due to space limitation.

disambiguated by either accent placement or prosodic phrasing (see examples (4) and (5)). She found that the participants performed well in L2 when the meaning contrast was realised prosodically in a similar way in their L1 and L2 (due to positive transfer) or in a language-independent way.

(4) She gave her dog biscuits.
(5) She dressed and fed the baby.

Atoye (2005) extended the study by Cruz-Ferreira (1989) and tested both discrimination of contours (perception) and interpretation in Nigerian learners of English using a sub-set of the stimuli from Cruz-Ferreira (1989). He found that the Nigerian learners were generally able to perceive differences between two prosodic versions of a pair but had substantial difficulties in glossing their meanings. These results suggest adequate low-level perceptual skills but difficulties in establishing the form-function link, similar to findings from Baker (2010).

Regarding the perception of paralinguistic meaning, A. Chen (2009) investigated the perception of gradient form-function mapping between pitch (scaling and alignment) and the attributes 'emphatic' and 'surprised' in Dutch learners of English and English learners of Dutch. The learners show a transfer from their L1, but it was also evident that they partially interpret non-L1-like form-function mappings in a native-like manner, due to language-independent uses of pitch and exposure to native input.

44.4.2 Perceived foreign accent and ease of understanding

When assessing L2 speech we can distinguish between perceived degree of foreign accent (or the linguistic nativelikeness) and ease of understanding (Derwing and Munro 1997; Munro and Derwing 1995). The latter is usually divided into intelligibility (number of words actually understood by listeners) and comprehensibility (how well listeners think they understand the speaker). Past work has shown that prosodic properties like speech rate, rhythm, intonation, and fluency can not only have an impact on accentedness ratings, but also on intelligibility and comprehensibility ratings. However, as different prosodic properties and aspects of L2 speech were examined in different L2s, we do not yet have a clear understanding of the effects of different prosodic properties on accentedness, intelligibility and comprehensibility in different L1-L2 parings. For example, Jilka (2000) found that for German learners of English accent

ratings were best predicted by f0-range and word stress measures whereas comprehensibility scores were mostly associated with speaking rates. Polyanskaya, Ordin and Busa (2017) found that for French learners of English both speech rate and speech rhythm (operationalized as durational ratios of syllables, vocalic sequences and consonantal clusters) influenced the degree of perceived foreign accent, but the effect of speech rhythm was larger than that of speech rate. Maastricht, Krahmer and Swerts (2016) showed that L2 Dutch containing deviance in pitch accent distributions for the purpose of focus marking produced by Spanish learners was rated as more foreign accented and more difficult to understand than L1 Dutch, with speakers' proficiency as a modulating factor. Using a cross-modal priming paradigm, Braun, Dainora and Ernestus (2011) showed that an unfamiliar intonation contour on otherwise Dutch sentences resulted in longer lexical decision latencies and semantic category judgments, suggesting an effect of non-native intonation on comprehensibility.

Much research on the assessment of L2 speech has been devoted to the issue of the weighting between segmental and prosodic characteristics in their impact on accentedness and comprehensibility. The findings have been rather mixed. Some studies found that deviations on the segmental level are less severe for the ratings of accentedness and comprehensibility than deviations on the prosodic level (Munro and Derwing 1995 on Mandarin learners of English; Trofimovich and Baker 2006 on Korean learners of English). On the other hand, recent studies have shown that segmental errors and the interplay between segments and prosody have a larger impact. For example, in an investigation of German-accented English by Ulbrich and Mennen (2016) the native listeners were more influenced by segments than prosody. In addition, listeners were quite sensitive to small prosodic differences when mixed with non-native segments. In a study with Korean-accented English, Sereno, Lammers and Jongman (2016) show that segments had a significant effect on accentedness, comprehensibility, and intelligibility, but intonation only had an effect on intelligibility.

Because different studies focused on different L2s produced by speakers with different L1s, future research is needed to better understand whether the weighting between segmental and prosodic properties in perceived accentedness and comprehensibility varies between L2 produced by learners with different L1s and between different L2s produced by learners with the same L1. We can state that both, segmental and prosodic features are responsible, differing in weighing from case to case. Thus, we can assume that an L2 learner who concentrates on segmental acuity leaving out prosodic aspects will probably be perceived as less fluent and less intelligible than an L2 speaker who cares more for fluency and prosody than for vowels and consonants.

44.5 Conclusions

Research on L2 sentence prosody covers many different aspects (intonation, timing) from different perspectives (production, perception, comprehension). Studies that integrate different aspects and perspectives are rare (but see Gut 2009a; Maastricht et al. 2016; Maastricht, Zee, Krahmer and Swerts 2017). For this reason theoretical models, which aim to predict learning difficulties only cover sub-fields of L2 sentence prosody. One example is L2 Intonational Learning Theory by Mennen (2015), which describes development in L2 intonational production along the same dimensions in which cross-linguistic differences in intonation can occur (Ladd 1996, 2008), i.e. the inventory of phonological prosodic elements, their distribution, how these elements are phonetically implemented, which functions these elements have, and how often these elements are used. Models that aim to explain the underlying learning mechanisms (e.g. answering questions like "What drives the learning of L2 prosody in the absence of explicit instructions?" or "What factors matter to the successful learning of L2 prosody?") are still lacking, even though much theoretical advances have been made in the acquisition of non-prosodic aspects of L2.

Existing research on L2 sentence prosody has been focused on the influence of L1 on L2 prosody. However, a solid analysis of such a transfer often encounters difficulties due to an unclear reference of 'correct' forms in the sentence prosody of the target variety. In contrast to L2 word prosody and L2 segmental forms, we frequently have greater optionality in L2 sentence prosody (e.g. in placing phrase breaks and pitch accents). We thus face a huge variability that is often enforced by regional and other non-standard influences. In addition, most studies are concerned with a variety of English as the target language. For L2 sentence prosody research it remains a challenge to define what is ‘correct’ or 'acceptable' in the target variety on the one hand, and to widen our knowledge of target and source languages on the other.

But there are more challenges for future research. Specifically, characteristics that are traditionally considered segmental, such as reduction, should be examined from a prosodic perspective. There is a need of investigating the interface between prosody and other linguistic levels such as syntax in L2 (e.g., Zubizarreta and Nava 2011). It is also important to extend L2 prosody research to more natural situations such as in dialogues and other interactional behaviour (e.g., Ward and Gallardo 2017). For this purpose, phonetic learner corpora with prosodic annotation would be a valuable source. At the same time learner corpora can help to

test theoretical questions about L2 sentence prosody with a substantial number of participants. This will in turn allow us to gain more insight into the developmental stages of L2 sentence prosody, ideally by establishing a hierarchy of learning difficulties. Finally, fluency and timing are not treated together with intonation and pitch-related aspects in L2 teaching, L2 assessment and L2 testing. However, such a paramount view on L2 sentence prosody would be beneficial to constructions of theories for acquisition of L2 prosody and applications such as assessments in teaching, exercises for individual learning and automatic testing of spoken performances.

Acknowledgements

We are grateful to Elisabeth Delais-Roussarie, Christiane Ulbrich, Katharina Zahner, Laura Downing, Aoju Chen and Carlos Gussenhoven for comments on an earlier version of this chapter. We also thank Clara Huttenlauch, Sophie Kutscheid, Jana Neitsch, Katharina Zahner, Maximilian Wiesner for help with references.

References

Abercrombie, David (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Akker, Evelien and Anne Cutler (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition 6*(2): 81-96.

Albin, Aaron Lee (2015). *Typologizing native language influence on intonation in a second language: Three transfer phenomena in Japanese EFL learners.* Unpublished PhD thesis, Indiana University.

Aoyama, Katsura and Susan G. Guion (2007). Prosody in second language acquisition. Acoustic analyses of duration and F0 range. In Ocke-Schwen Bohn and Murray J. Munro (eds.), *Language Experience in Second Language Speech Learning: In Honour of James Emil Flege*. Amsterdam: Benjamins, 281-297

Arnhold, Anja (2016). Complex prosodic focus marking in Finnish: expanding the data landscape. *Journal of Phonetics 56*: 85–109.

Arvaniti, Amalia, D. Robert Ladd and Ineke Mennen (2006). Tonal association and tonal alignment: Evidence from Greek polar questions and contrastive statements. *Language and Speech 49*(4): 421-450.

Atoye, Raphael O. (2005). Non-native perception and interpretation of English intonation. *Nordic Journal of African Studies 14(1)*: 26-42.

Atterer, Manuela and D. Robert Ladd (2004). On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics 32*(3): 177-197.

Baker, Rachel Elizabeth (2010). *The Acquisition of English Focus Marking by Non-native Speakers.* Unpublished PhD thesis, Northwestern University.

Barry, William J. and Bistra Andreeva (2001). Cross-language similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic Association 31*: 51-66.

Bartels, Christine (1999). *The Intonation of English Statements and Questions. A Compositional Interpretation*. New York & London: Garland Publishing.

Bartkova, Katarina, Elisabeth Delais-Roussarie and Fabián Santiago Vargas (2012).

PROSOTRAN: A tool to annotate prosodically non-standard data. Paper presented at the *6th International Conference on Speech Prosody*. Shanghai, China.
Bosker, Hans R., Hugo Quené, Ted J. M. Sanders and Nivja H. de Jong (2014). Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language 75*: 104-116.
Braun, Bettina, Audra Dainora and Mirjam Ernestus (2011). An unfamiliar intonation contour slows down on-line speech comprehension. *Language and Cognitive Processes 26*: 350-375.
Braun, Bettina, Nicole Dehé, Jana Neitsch, Daniela Wochner and Katharina Zahner (2018). The prosody of rhetorical and information-seeking questions in German. *Language and Speech advance online publication*.
Braun, Bettina and Lara Tagliapietra (2011). On-line interpretation of intonational meaning in L2. *Language and Cognitive Processes 26*: 224-235.
Brown, Gillian (1983). Prosodic structure and the given/new distinction. In Anne Cutler and D. Robert Ladd (eds.), *Prosody: Models and Measurements*. Berlin: Springer, 67-77
Büring, Daniel (2009). Towards a typology of focus realizations. In Malte Zimmermann and Caroline Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives*. New York: Oxford University Press, 177-205
Butcher, Andrew (1981). Aspects of the speech pause: phonetic correlates and communicative functions. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK) 15*.
Calbert Graham, Brechtje Post (2018). Second language acquisition of intonation: Peak alignment in American English. *Journal of Phonetics 66*: 1-14.
Chen, Aoju (2009). Perception of paralinguistic intonational meaning in a second language. *Language Learning 59(2)*: 367 - 409.
Chen, Aoju and Diantha De Jong (2015). Prosodic expression of sarcasm in L2 English. Spoken discourse and prosody in L2. *Series Materiali Linguistici*.
Chen, Y., Y. Xu, S. Guion-Anderson (2014). Prosodic Realization of Focus in Bilingual Production of Southern Min and Mandarin. *Phonetica 71(4)*: 249-270.
Chen, Yang, Michael P. Robb, Harvey R. Gilbert and Jay W. Lerman (2001). A study of sentence stress production in Mandarin speakers in American English. *Journal of the Acoustical Society of America 109*(4): 1681-1690.
Corley, Martin, Lucy J. MacGregor and David I. Donaldson (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition 105*: 658-668.
Cruz-Ferreira, Madalena (1989). *Non-Native Comprehension of Intonation Patterns in Portuguese and in English.* University of Manchester.
Crystal, David (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
Cucchiarini, Catia, Helmer Strik and Lou Boves (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America 111*(6): 2862-2873.
de Jong, Nivja H. (2016). Fluency in second language assessment. In Dina Tsagari and Jayanti Banerjee (eds.), *Handbook of Second Language Assessment*: Mouton de Gruyter, 203-218
Dehé, Nicole and Bettina Braun (2013). The prosody of question tags in English. *English Language and Linguistics 17*(1): 129-156.
Dehé, Nicole, Ingo Feldhausen and Shinichiro Ishihara (2011). The prosody-syntax interface: Focus phrasing, language evoluation. *Lingua 121*(13): 1863-1869.
Delais-Roussarie, Elisabeth and Hi-Yon Yoo (2011). Learner corpora and prosody: From the COREIL corpus to principles on data collection and corpus design. *Poznań Studies in*

Contemporary Linguistics 47(1): 26-39.
Derwing, Tracey M. and Murray J. Munro (1997). Accent, intelligibility, and comprehensibility. Evidence from four L1s. *Studies in Second Language Acquisition 20*: 1-16.
Domaneschi, Filippo, Maribel Romero and Bettina Braun (2017). Bias in polar questions. *Glossa: a journal of general linguistics 21*(2): 23.
Doughty, Catherine J. and Michael H. Long (Eds.). (2005). *The Handbook of Second Language Acquisition*. Malden: Wiley-Blackwell.
Downing, Laura J. and Annie Rialland (2017). *Intonation in African Tone Languages*: Berlin/Boston: Walter de Gruyter.
Duez, Danielle (1982). Silent and non silent pauses in three speech styles. *Language and Speech 25*: 11-28.
Eady, Stephen J., William E. Cooper, Gayle V. Klouda, Pamela R. Mueller and Dan W. Lotts (1986). Acoustical characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments. *Language and Speech 29*(3): 233-251.
Engstrand, Olle and Diana Krull (2001). Simplification of phonotactic structures in unscripted Swedish. *Journal of the International Phonetic Association 31*: 41-50.
Gass, Susan M. and Alison Mackey (Eds.). (2012). *The Routledge Handbook of Second Language Acquisition*. London & New York: Routledge.
Götz, Sandra (2013). *Fluency in native and nonnative English speech* Amsterdam: John Benjamins.
Grabe, Esther (1998). Pitch accent realization in English and German. *Journal of Phonetics 26*: 129-143.
Grosjean, François and Maryann Collins (1979). Breathing, pausing and reading. *Phonetica 36*(2): 98-114.
Gut, Ulrike (2009a). *Introduction to English Phonetics and Phonology*. Frankfurt: Peter Lang.
Gut, Ulrike (2009b). Non-native speech. A corpus-based analysis of phonological and phonetic properties of L2 English and German. Frankfurt: Peter Lang,
Gut, Ulrike (2012). The LeaP corpus A multilingual corpus of spoken. In Thomas Schmidt and Kai Wörner (eds.), *Multilingual corpora and multilingual corpus analysis. Hamburg Studies on Multilinguialism 14*. Amsterdam: John Benjamins, 3-23
Gut, Ulrike and Stefanie Pillai (2014). Prosodic marking of information structure by Malaysian speakers of English. *Studies in Second Language Acquisition 36*(2): 283-302.
Hedberg, Nancy, Juan M. Sosa and Lorna Fadden (2004). Meanings and configurations of questions in English. Paper presented at the *Speech Prosody*. Nara, Japan.
Hedberg, Nancy, Juan M. Sosa, Emrah Gürgülü and Morgan Mameni (2010). Prosody and pragmatics of wh-interrogatives. Paper presented at the *2010 Meeting of the Canadian Linguistics Association*.
Herment, Sophie, Nicolas Ballier, Elisabeth Delais-Roussarie and Anne Tortel (2014). Modelling interlanguage intonation: the case of questions. Paper presented at the *7th International Conference on Speech Prosody*. Dublin, Ireland.
Herment, Sophie, Anastasia Loukina and Anne Tortel. (2012). The AixOx corpus. http://sldr.org/sldr000784/fr (last accessed June 2017).
Herschensohn, Julia and Martha Young-Scholten (Eds.). (2013). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press.
Hertel, Tammy Jandrey (2003). Lexical and discourse factors in the second language acquisition of Spanish word order. *Second Language Research 19*(4): 273-304.
Horgues, Celine (2013). French learners of L2 English: Intonation boundaries and the

marking of lexical stress. *Research in Language 11*(1): 41-56.
Huang, Becky H. and Sun-Ah Jun (2011). The effect of age on the acquisition of second language prosody. *Language and Speech 54*(Pt 3): 387-414.
Jilka, Matthias (2000). *The Contribution of Intonation to the Perception of Foreign Accent.* University of Stutgart, PhD Thesis.
Jun, S.-A. and Mira Oh (2000). Aquisition of second language intonation. Paper presented at the *Proceedings of International Conference on Spoken Language Processing*. Beijing, China.
Jun, Sun-Ah (2014). Prosodic typology: by prominence type, word prosody, and macro-rhythm. In Sun-Ah Jun (ed.), *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 520-539
Jun, Sun-Ah (Ed.). (2005). *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press.
Kainada, Evia and Angelos Lengeris (2015). Native language influence on the production of second-language prosody. *Journal of the International Phonetic Association 45*(3): 269-287.
Kelm, Orlando R. (1987). An acoustic study on the differences of contrastive emphasis between native and nonnative Spanish. *Hispania 70*(3): 627-633.
Kohler, Klaus J. (2004). Pragmatic and attitudinal meanings of pitch patterns in German syntactically marked questions. In Gunnar Fant, Hiroya Fujisaki, J. Cao, et al. (eds.), *From traditional phonology to modern speech processing - Festschrift für Professor Wu Zongji's 95th Birthday*. Bejing: Foreign Language Teaching and Research Press, 205-215
Krahmer, E. and M. Swerts (2001). On the alleged existence of contrastive accents. *Speech Communication 34*: 391-405.
Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica 55*: 243-276.
Kula, N. C. and S. Hamann (2017). Intonation in Bemba. *Intonation in African Tone Languages 24*: 321-363.
Ladd, D. Robert (1980). *The Structure of Intonational Meaning: Evidence from English*. Bloomington: Indiana Universtiy Press.
Ladd, D. Robert (1996). *Intonational Phonology*. Cambridge [England]; New York, NY, USA: Cambridge University Press.
Ladd, D. Robert (2008). *Intonational Phonology* (2nd ed. Vol. 119). Cambridge: Cambridge University Press.
Ladefoged, Peter (1975). *A course in phonetics*. New York: Harcourt Brace Jovanovich
Lee, Eun-Kyung and Scott H. Fraundorf (2016). Effects of contrastive accents in memory for L2 discourse. *Bilingualism: Language and Cognition ahead of print*: 1-17.
Levelt, Willem J.M. (1983). Monitoring and self-repair in speech. *Cognition 14*: 41-104.
Li, Aike and Brechtje Post (2014). L2 Acquisition of prosodic properties of speech rhythm: Evidence from L1 Mandarin and German learners of English. *Studies in Second Language Acquisition 36*(2): 223-255.
Liang, J. and Vincent Van Heuven, J. (2007). Chinese tone and intonation perceived by L1 and L2 listeners. In Carlos Gussenhoven and Thomas Riad (eds.), *Tones and Tunes, Volume 2: Experimental studies in word and sentence prosody*. Berlin/New York: Mouton de Gruyter, 27-61
Maastricht, Lieke J. van, Emiel Krahmer and Marc Swerts (2016). Native speaker perceptions of (non-)native prominence patterns: Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication 83*: 21-33.

Maastricht, Lieke J. van, Tim Zee, E.J. Krahmer and M.G.J. Swerts (2017). L1 perceptions of L2 prosody: The interplay between intonation, rhythm, and speech rate and their contribution to accentedness and comprehensibility. Paper presented at the *Proceedings of Interspeech*. Stockholm.

McGory, Julie (1997). *The acquisition of intonation patterns in English by native speaker of Korean and Mandarin.* Unpublished PhD thesis, Ohio State University, Columbus.

Mennen, Ineke (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics 32*: 543-563.

Mennen, Ineke (2015). Beyond segments: Towards an L2 intonation learning theory. In Elisabeth Delais-Roussarie, Mathieu Avanzi and Sophie Herment (eds.), *Prosody and Language in Contact - L2 Acquisition, Attition and Languages in Multilingual Situations*. Berlin: Springer, 171-188

Mennen, Ineke and E. de Leeuw (2014). Beyond segments: prosody in SLA. *Studies in Second Language Acquisition 36*(2): 183-194.

Munro, Murray J. and Tracey M. Derwing (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning 45*: 73–97.

Nagano-Madsen, Y. (2014). Acquisition of L2 Japanese intonation - data from Swedish learners (in Japanese). *Japanese Speech Communication 2*: 1-27.

Nespor, Marina and Irene Vogel (2007). *Prosodic Phonology* (2nd edition ed.). Berlin: de Gruyter.

Nguyễn, T. Anh-Thur, C.L. John Ingram and J. Rob Pensalfini (2008). Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics 36*(1): 158-190.

Nickels, Stefanie, Bertram Opitz and Karsten Steinhauer (2013). ERPs show that classroom-instructed late second language learneres rely on the same prosodic cues in syntactic parsing as native speakers. *Neuroscience Letters 557*: 107-111.

Nickels, Stefanie and Karsten Steinhauer (2018). Prosody-syntax integration in a second language: Contrasting event-related potentials from German and Chinese learners of English using linear mixed effect models. *Second Language Research 34*(1): 9-37.

O'Brien, Mary Grantham and Ulrike Gut (2010). Phonological and phonetic realization of different types of focus in L2 speech. In K. Dziubalska-Kołaczyk, M. Wrembel and M. Kul (eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds*. Frankfurt: Peter Lang, 205-215

O'Connor, Joseph Desmond and Gordon Frederick Arnold (1973). *Intonation of Colloquial English*. London: Longman.

Ordin, Mikhail and Leona Polyanskaya (2015). Perception of speech rhythm in second language: the case of rhythmically similar L1 and L2. *Frontiers in Psychology 6*: 316.

Ortega-Llebaria, Marta and Laura Colantoni (2014). L2 English intonation. Relations between form-meaning associations, access to meaning and L1 transfer. *Studies in Second Language Acquisition 36*(2): 331-353.

Patil U, Kentner G, Gollrad A, Kügler F, Fery C, Vasishth S (2008). Focus, word order and intonation in Hindi. *J South Asian Linguist 1*: 55–72.

Pierrehumbert, Janet B. and Julia Hirschberg (1990). The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan and Martha E. Pollack (eds.), *Intentions in Communication*. Cambridge: MIT Press, 271-311

Pike, Kenneth L. (1945). *The Intonation of American English*. Michigan: University of Michigan Publications.

Polyanskaya, Leona, Mikhail Ordin and M.G. Busa (2017). Relative salience of speech rhythm and speech rate on perceived foreign accent in L2. *Language and Speech 60*:

333-355.
Puga, Karin, Robert Fuchs, Jane Setter and Peggy Mok (2017). The Perception of English Intonation Patterns by German L2 speakers of English. Paper presented at the *Interspeech 2017*. Stockholm, Sweden.
Pürschel, Heiner (1975). *Pause und Kadenz. Interferenzerscheinungen bei der englischen Intonation deutscher Sprecher*. Tübingen: Niemeyer.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (Eds.). (1985). *A Comprehensive Grammar of the English Language* (Vol. Longman): New York.
Ramírez Verdugo, Dolores (2005). The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty. *Journal of Pragmatics 37*(12): 2086-2115.
Ramírez Verdugo, Dolores (2006). Prosodic realization of focus in the discourse of Spanish learners and English native speakers. *Estudios Ingleses de la Universidad Complutense 14*: 9-32.
Ramírez Verdugo, Dolores and J. Romero Trillo (2005). The pragmatic function of intonation in L2 discourse: English tag questions used by Spanish speakers. *Intercultural Pragmatics 2*(2): 151-168.
Rasier, Laurent and Philippe Hiligsmann (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique Française 28*: 41-66.
Raupach, Manfred (1980). Temporal variables in first and second language speech production. In Hans W. Dechert and Manfred Raupach (eds.), *Temporal Variables in Speech*. The Hague: Mouton de Gruyter, 263-270
Ritchie, William and Tej K. Bhatia (Eds.). (2009). *The New Handbook of Second Language Acquisition*. Bingley: Emerald.
Saito, Hiroko (2006). Nuclear-stress placement by Japanese learners of English: Transfer from Japanese. In Y. Kawaguchi, I. Fónagy and T. Moriguchi (eds.), *Prosody and syntax*. Amsterdam: John Benjamins, 125-139
Santiago Vargas, Fabián and Elisabeth Delais-Roussarie (2012). Acquiring phrasing and intonation in French as second language: the case of yes-no questions produced by Mexican Spanish learners. Paper presented at the *6th International Conference on Speech Prosody*. Shanghai, China.
Schwab, Sandra and François Grosjean (2004). La perception du débit en langue seconde [The perception of speech rate in a second language]. *Phonetica 61*: 84-94.
Sereno, Joan, Lynne Lammers and Allard Jongman (2016). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics 37*(2): 303-322.
Smorenburg, Laura, Joe Rodd and Aoju Chen (2015). The effect of explicit training on the prosodic production of L2 sarcasm by Dutch learners of English. Paper presented at the *18th International Congress of the Phonetic Sciences*. Glasgow, UK.
Steindel Burdin, Rachel, Sara Phillips-Bourras, Rory Turnbull, Murat Yasavul, Cynthia G. Clopper and Judith Tonhauser (2014). Variation in the prosody of focus in head- and head/edge-prominence languages. *Lingua 165*: 254-276.
Swerts, Marc and Sabine Zerbian (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica 67*: 127-146.
Trofimovich, Pavel and Wendy Baker (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.
Trouvain, Jürgen (2004). *Tempo variation in speech production.* Saarland University, PhD Thesis.
Trouvain, Jürgen, Camille Fauth and Bernd Möbius (2016). Breath and non-breath pauses in

fluent and disfluent phases of German and French L1 and L2 read speech. Paper presented at the *Proceedings of the 8th International Conference on Speech Prosody*. Boston.
Trouvain, Jürgen and Martine Grice (1999). The effect of tempo on prosodic structure. Paper presented at the *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*. San Francisco.
Trouvain, Jürgen and Bernd Möbius (2014). Sources of variation of articulation rate in native and non-native speech: comparisons of French and German. Paper presented at the *Proceedings of the 7th International Conference on Speech Prosody*. Dublin.
Trouvain, Jürgen, Frank Zimmerer, Bernd Möbius, Mária Gósy and Anne Bonneau (2017). Segmental, prosodic and fluency features in phonetic learner corpora - Introduction to special issue. *International Journal of Learner Corpus Research 3*(2): 105-118.
Truckenbrodt, Hubert (2012). Semantics of intonation. In Claudia Maienborn, Klaus von Heusinger and Paul Portner (eds.), *Semantics. An International Handbook of Natural Language Meaning*. Berlin: de Gruyter, 2039-2069
Turco, Giuseppina, Christine Dimroth and Bettina Braun (2015). Prosodic and lexical marking of contrast in L2 Italian. *Second Language Research 31*: 465-491.
Ueyama, Motoko and Sun-Ah Jun (1998). Focus realization in Japanese English and Korean English intonation. *Japanese and Korean Linguistics 7*: 629-645.
Ulbrich, Christiane (2013). German pitches in English: Production and perception of cross-varietal differences in L2. *Bilingualism: Language and Cognition 16*(2): 397-419.
Ulbrich, Christiane and Ineke Mennen (2016). When prosody kicks in: The intricate interplay between segments and prosody in perceptions of foreign accent. *International Journal of Bilingualism 20*(5): 522-549.
Vainio, Martti and Juhani Järvikivi (2007). Focus in production: tonal shape, intensity and word order. *Journal of the Acoustical Society of America 121*: EL55–EL61.
Vallduví, Enric and Elisabet Engdahl (1996). The linguistic realisation of information packaging. *Linguistics 34*: 459-519.
Ward, Nigel G. and Paola Gallardo (2017). Non-native differences in prosodic-construction use. *Dialogue and Discourse 8*(1): 1-30.
Wells, John C. (2006). *English intonation: An introduction*. Cambridge: Cambridge University Press.
Wennerstrom, Ann (1994). Intonational meaning in English discourse: A study of non-native speakers. *Applied Linguistics 15(4)*: 399 – 420.
White, Laurence and Sven L. Mattys (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics 35*: 501-522.
Wieden, Wilfried (1993). Aspects of acquisitional stages. In Bernhard Kettemann and Wilfried Wieden (eds.), *Current Issues in European Second Language Acquisition Research*. Tübingen: Gunter Narr Verlag, 125-135
Wiese, Richard (1983). *Psycholinguistische Aspekte der Sprachproduktion*. Hamburg: Buske.
Xu, Yi (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics 27*: 55-105.
Xu, Yi, Szu-Wei Chen and Bei Wang (2012). Prosodic focus with and without post-focus compression (PFC): a typological divide within the same language family? *Linguist Review 29*: 131–147.
Yang, Chunsheng and Marjorie K.M. Chan (2010). The perception of Mandarin Chinese tones and intonation by American learners. *Journal of the Chinese Teachers Association 45*(1): 7-36.
Zahner, Katharina and Jenny Yu (2019). Compensation strategies in non-native English and German. Paper presented at the *ICPhS 2019*. Melbourne, Australia.

Zimmerer, Frank, Bistra Andreeva, Jeanin Jügler and Bernd Möbius (2015). Comparison of pitch profiles of German and French speakers speaking French and German. Paper presented at the *18th International Congress of Phonetic Sciences*. Glasgow, UK.
Zubizarreta, Maria Luisa and Emily Nava (2011). Encoding discourse-base meaning: Prosody vs. Syntax. Implications for Second Language Acquisition. *Lingua 121*: 652-669.