Studies in Second Language Acquisition, 2014, **36**, 331–353.
doi:10.1017/S0272263114000011

L2 ENGLISH INTONATION

Relations between Form-Meaning Associations, Access to Meaning, and L1 Transfer

Marta Ortega-Llebaria
University of Pittsburgh

Laura Colantoni
University of Toronto

Although there is consistent evidence that higher levels of processing, such as learning the form-meaning associations specific to the second language (L2), are a source of difficulty in acquiring L2 speech, no study has addressed how these levels interact in shaping L2 perception and production of intonation. We examine the hypothesis of whether access to contextual meaning increases the chances of first language (L1) influence on L2 intonation. To test this hypothesis, we compared the perception and production of sentential English focus by 27 advanced English language learners (n = 13 L1 Mandarin speakers; n = 14 L1 Spanish speakers) and 13 controls, through a series of tasks that promoted different levels of access to meaning. Results showed that L1 transfer was especially clear in Spanish speakers. Not only did they consistently differ from controls in their perception

This research has been supported by a SSHRC grant (890-2011-0049) to both authors, a Victoria College research grant to Laura Colantoni, and a research grant from Arts and Sciences College, University of Pittsburgh, to Marta Ortega-Llebaria. The authors want to thank Charlotte Rogers, Julia Sandoval, Anita Rao, Lexi Williams, Olivia Marasco, and Matt Patience for their assistance during the creation of the materials, subject testing, and data labeling.

Correspondence concerning this article should be addressed to Marta Ortega-Llebaria, University of Pittsburgh, Department of Linguistics, 2830 Cathedral of Learning, 4200 Fifth Avenue, Pittsburgh, PA 15260. E-mail: mao61@pitt.edu

of focalized verbs and subjects, showing their L1 bias to perceive focus at the end of a sentence, but they were also the only group of speakers that inserted pauses after the focalized word, which showed strong L1 effects. Moreover, these L1 transfer effects were more obvious in contextualized tasks, which indicated that facilitating access to meaning by adding context increased L1 transfer effects on the perception and especially on the production of focus intonation.

BACKGROUND

Perception of L2 Segments and Stress

It is well established that exposure to one's native language molds speech perception. By 8 months of age, infants have changed from a universal to a language-specific mode of perception by paying more attention to the regularities of their native language. For example, Werker and Tees (1984, 2002) showed that four 1-month-old English and Hindi infants discriminated equally well two Hindi sounds that did not exist in English. At 8 months of age, however, only Hindi infants maintained a high discrimination of the contrast, whereas English infants' discrimination decreased, approximating that of native English-speaking adults. This attunement to the regularities of one's own native language, which increases progressively with language exposure, leads to the formation of language-specific categories. As a result, adults who are learning a second language (L2) process L2 speech through the regularities and categories of their first language (L1), which may hinder L2 perception. The degree of difficulty posed by L1 transfer to L2 speech perception varies according to, among other factors, the degree of similarity between the speaker's L1 and the L2 being learned, as shown by models of L2 speech (e.g., Best, 1994; Best & Tyler, 2007; Flege, 1995). Moreover, these L1 effects on adult L2 speech perception can recede with sufficient L2 input (Flege & Frieda, 1997; Piske, MacKay, & Flege, 2001).

Given the wide consensus regarding the modeling of speech perception on the basis of language exposure, recent research has centered on which levels of processing are affected by such exposure. On the one hand, there is supporting evidence for the view that auditory processing of speech sounds remains unchanged by language experience and that the observed changes in speech perception relate to higher level processes, such as phonological encoding (Best & Tyler, 2007) or speech-specific selective attention (Pisoni, Lively, & Logan, 1994). For example, adults are able to discriminate nonnative sounds that they cannot categorize linguistically (Werker & Tees, 1984). Moreover, Dupoux and colleagues

(Dupoux, Pallier, Sebastián-Gallés, & Mehler, 1997; Dupoux, Peperkamp, & Sebastián-Gallés, 2001, 2010; Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008) showed that French adults could perceive lexical stress in Spanish in AX discrimination tasks with a short interstimulus interval, which promote acoustic perception. However, the same French adults showed "stress deafness" in tasks tapping phonological encoding. On the other hand, neurophysiological studies have shown that language-specific sensitivities are present in mismatch-negativity event-related potentials, suggesting that perceptual changes due to language exposure may also occur at early phonetic or late auditory levels, prior to the recognition of speech in terms of higher level linguistic units (Nenonen, Shestakova, Huotilainen, & Näätänen, 2003, 2005).

Perception of L2 Sentence Intonation

The preceding discussion about which levels of processing are affected by language exposure is mainly based on the perception of L2 segments and stress; research on L2 sentence intonation has not paid particular attention to the effects of different levels of processing. For example, Cruz-Ferreira (1984) compared L1 English-L2 Portuguese and L1 Portuguese-L2 English speakers' capacity to match a sentence with a specific intonation with one of three alternative glosses; the results showed consistent differences between natives and L2 learners. Chen, Rietveld, and Gussenhoven (2001) tested crosslinguistic (English vs. Dutch) differences in the perception of a pragmatic continuum (confident vs. nonconfident and friendly vs. nonfriendly) by manipulating the pitch range and documented some language-specific and universal trends in the perception of intonation, such as bias toward perceiving English as more friendly than Dutch at identical pitch ranges.

To our knowledge, only Grabe, Rosner, García-Albea, and Zhou (2003) have directly addressed the issue of processing levels and their shaping of perception of L2 sentence intonation by examining subjects' L2 perception on tasks aimed at different levels of processing. More specifically, they examined the perception of the phrase *Melany Maloney* spoken with 11 speech contours common in southern British English by native English, Chinese, and Spanish speakers. A similar group of speakers listened to an identical set of the 11 fundamental frequency (F0) contours on frequency-modulated sine waves. In both tasks, listeners had to decide if stimuli were the same or different. Similarity judgments on the first task were based on a higher level of processing than those on the second task. Results showed L1 effects only in the first task, which led the authors to conclude that "the perception of similarities and differences

among intonation contours calls upon universal auditory mechanisms whose output is molded by experience with one native language" (Grabe et al., 2003, p. 375).

Thus, on the one hand, Grabe and colleagues' (2003) study shows that L1 experience molds the perception of L2 sentence intonation as soon as there is a possible association between the F0 contour of an isolated phrase, *Melanie Maloney*, and a pragmatic/semantic interpretation, such as statement, question, or surprise. On the other hand, studies such as Cruz-Ferreira (1989), in which sentences are contextualized to precisely define meaning, thereby making meaning more accessible to L2 learners, still show L1 transfer effects in the association of L2 meaning to a L2 sentence-level F0 contour. Altogether, these studies suggest that L1 influence on L2 form-meaning associations is present through different processing levels that imply different degrees of access to meaning. No study, however, has addressed the issue of how different levels of access to meaning shape the perception (and production) of L2 sentence intonation.

To fill in this gap, we examine whether L1 transfer on L2 form-meaning associations can be modeled by manipulating access to meaning through context (see Figure 1). If this modeling occurs, there are two logical scenarios; namely, facilitation of access to meaning either reduces L1 transfer or increases it. On the basis of previous research that shows that tasks that promoted acoustic perception triggered less L1 transfer than those involving access to meaning or higher levels of processing, we hypothesize that increasing access to meaning will increase L1 transfer effects.

To test this hypothesis, we compared the perception of English contrastive sentence focus in subject, verb, and object position by advanced learners of English whose L1 was Spanish or Mandarin Chinese to that of an English control group. Two perception and two production tasks were administered to the three groups of speakers. Access to

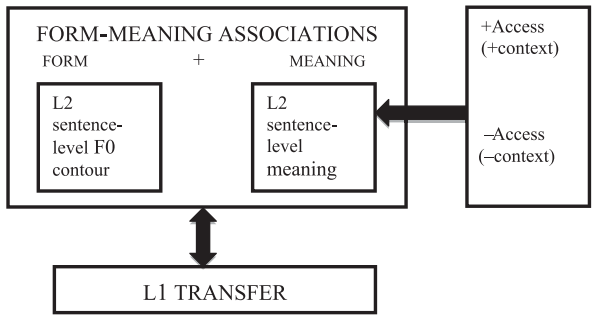

Figure 1. Relations among sentence intonation, access to meaning, and L1 transfer.

meaning in the perception and production tasks was manipulated by controlling context. In the +access-to-meaning tasks (+AM), participants had to produce or perceive sentences related to a story (with illustrations) that they listened to beforehand. Thus, these +AM tasks required participants to match a contextualized and accessible meaning to the appropriate sentence F0 contour. In the –access-to-meaning tasks (–AM), participants were presented with isolated sentences unrelated in meaning. In the production task, they had to imitate these sentences, whereas in the perception task, they had to match the sentence with the intonation contour of one out of three low-pass filtered sentences. These –AM tasks prompted participants to choose a sentence F0 contour without a clear and accessible meaning. In fact, it was possible for the participants to perform the task by paying attention mainly to the F0 contour while disregarding any precise meaning related to prominence.

First language transfer was controlled by testing participants from two L1s that differ in their language-specific strategies to express contrastive meaning. These language-specific differences, together with a brief review of previous literature on L2 English focus by Spanish and Chinese speakers, are explained in the next two sections and will guide our predictions, which are specified following these two sections.

Sentence Focus in English, Mandarin Chinese, and Spanish

In English, a word is made prominent in comparison to other words in the sentence by two strategies. First, the prominent word may be associated with a pitch accent of expanded pitch range, longer duration, and louder intensity. Second, words following the prominent word are produced with compressed pitch range, shorter durations, and softer intensity. As a result, the prominent word stands out from the nonprominent words (i.e., words in the background; Bolinger, 1989; Cruttenden, 1997; Gussenhoven, 2004; Ladd, 2008). The meaning attributed to this F0 contour relates to whether a word is new or given in the discourse by highlighting new words while placing given words in the background (but see Watson, Arnold, & Tanenhaus, 2008). Another possible interpretation is that of contrastive focus, in which, in addition to being new, the highlighted word conveys a clarification meaning. For example, when answering *No, JOHN is coming tomorrow* to the question *Who is coming tomorrow? Mary?*, *John* is the new information: It clarifies that it is *John* and not *Mary* who is coming tomorrow, and it is spoken with an expanded pitch range, longer duration, and louder intensity than the words *is coming tomorrow*. This makes *John* the most prominent word in the sentence. English can make use of this intonation strategy to

mark focus in different sentence positions (e.g., a sentence such as *MARY is leaving tomorrow* may be a clarification of the statement *Peter is leaving tomorrow*, whereas *Mary is LEAVING tomorrow* may be a response to the question *Is Mary arriving tomorrow?*). Thus, English is relatively plastic in the use of prosodic cues to mark sentence prominence but has a relatively fixed word order (see Vallduví, 1991, for details on the proposal of [+/– plastic] languages, and Ladd, 2008, Chapter 6, for the different theoretical positions regarding the prosodic marking of focus).

Unlike English, there is disagreement on whether Spanish uses pitch accents to convey the meaning of contrastive focus. Whereas Face (2001) found evidence for pitch accents that signal different types of foci in Madrid Spanish, Toledo (1989) did not find clear acoustic cues in Buenos Aires Spanish. There is, however, complete agreement that, across dialects, Spanish has syntactic resources to mark prominence, and these involve changes in word order. As a result, a phrase with a postverbal subject, such as *Llega María* "is-arriving Maria" is an appropriate answer to the question *What is happening?* (broad focus), whereas *María llega* "Mary is-arriving," with a preverbal subject, receives a contrastive interpretation (Zubizarreta, 1998). Thus, whereas a contrastive meaning in English is expressed by intonation (*MARY is arriving*), an equivalent meaning in Spanish is expressed by word order (*María llega*).

In Mandarin, sentence prosody is constrained by lexical tone preservation (Y. Xu, 1999, 2001; Yip, 2002). For example, in Mandarin the word /ma/ may have four different meanings depending on whether it is produced with a high-level tone ("mother"), a high rising contour ("hemp"), a low falling-rising contour ("horse"), or a high falling tone ("scold"; e.g., Blicher, Diehl, & Cohen, 1990; Howie, 1976). Thus, preserving the pitch shapes in a sentence is essential for communication of lexical meaning. As a result, pragmatic meanings are expressed in ways that do not greatly distort the pitch shape of the lexical tones. This can be done (a) by expanding the pitch range of the tone in the focalized word (as well as the duration of the stressed syllable) without changing its shape or (b) by compressing the pitch range in postfocal words while maintaining the tones associated with those words (e.g., Bolinger, 1989; Liu, 2009; Wang & Xu, 2011). Prefocal words, however, are not affected by pitch compression (Liu, 2009). Finally, Mandarin, as English, has a default subject-verb-object (i.e., SVO) word order. As in English, contrastive focus can be prosodically marked in situ or by adding the contrastive focus particle *shi* (see L. Xu, 2004). Thus, whereas Spanish makes use of word order to express new versus given word status and contrastive focus meanings, equivalent meanings in Mandarin and English are expressed by intonation, by expanding the pitch range and duration of the prominent word, and by compressing the pitch range and duration of words in postfocal position.

Acquisition of English Sentence Focus by Native Speakers of Spanish and Chinese

To ground our predictions further, we briefly review the existing studies that have dealt with the acquisition of English intonation by either Spanish or Mandarin speakers. As concerns the acquisition of English intonation by Spanish speakers, very few details are known. Nava and Zubizarreta (2009) and Zubizarreta and Nava (2011) conducted experimental studies on the acquisition of English nuclear stress rules (i.e., the most prominent element in the intonation phrase) and anaphoric deaccenting (i.e., the deletion of a pitch accent on an element that was previously mentioned in the phrase) by a group of intermediate and advanced Spanish learners of English. They showed that the nuclear stress rule is acquired later than anaphoric deaccenting, in particular in SV sentences, for which Spanish learners tended to assign prominence to the verb (i.e., the last element in the phrase) instead of to the subject, which indicated a preference for the use of syntactic over prosodic means to mark sentence focus. These studies suggest that, in spite of the typological similarities between English and Spanish, some aspects of sentence prosody are not mastered even by advanced speakers.

More details are known about the acquisition of English intonation by Mandarin speakers. The most illustrative study for the present article is McGory (1997), which examines the acquisition of American English by Mandarin and Korean learners. A key piece of information for our study is that learners in McGory's study associated English stressed syllables with native rising tones. Thus, all lexical words in the utterance displayed a rising pitch accent with the peak aligned within the stressed syllable. This, in turn, had consequences for the realization of sentential focus, as evidenced by the fact that participants had difficulties with the production of prominence. In particular, Mandarin learners tended to assign equal prominence to all words in the sentence, as opposed to native English speakers, who used pitch accents only on prominent words.

Predictions

Given our hypothesis, we make the following predictions regarding L1 transfer (Predictions 1 and 2), access to meaning (Prediction 3), and production (Prediction 4).

1. Given that English and Mandarin use similar intonation strategies to express contrastive meaning, we predict positive transfer from L1 Mandarin to L2 English. The Mandarin L1 strategy will aid in correctly processing the form-meaning association of English sentence focus.

2. Given that Spanish expresses a contrastive meaning mainly by word order, not by intonation, we predict that Spanish speakers will experience difficulty when prominence is not on the final element (i.e., the object).
3. We expect that making meaning more accessible, which is assessed by comparing subjects' performance between the +AM and –AM tasks, will increase L1 transfer.
4. If perception leads to production, L1 transfer will be greater in production. Any effect of access to meaning on L1 transfer observed in perception will be observed as well in production.

METHOD

Participants

Three groups of participants were included in the study. Our controls were a group of 13 American English speakers (9 females, and 4 males), who were tested in Pittsburgh, Pennsylvania. The second group included 14 Spanish learners of English (10 females and 4 males), who were tested in Toronto, Ontario, Canada. The L1 Mandarin group was composed of 13 speakers (9 females and 4 males), who were tested either in Pittsburgh or in Toronto.

Participants were recruited on the basis of their age (18 or older), their L1 (Spanish, Mandarin, or English), and their knowledge of English. Although no specific level of English was targeted, participants were all living in an English-speaking country and were able to communicate in English, which was the language used to conduct the testing. Our initial recruiting included 25 L1 Mandarin and 21 L1 Spanish participants. Of those participants, we selected a subset on the basis of (a) their self-reported English proficiency (advanced and near-native speakers); (b) their overall performance in all tasks (the most fluent participants were selected); and (c) their relative similarity with respect to their age of onset of acquisition (AOA), length of residency (LOR) in either Canada or the United States, and number of years during which they had systematically studied English (see Table 1). Through a background questionnaire, we also collected information about participants' use of English, which is reported in Table 1.

Although an effort was made to find two similar experimental groups, Mandarin and Spanish learners of English differed in their mean age, in their AOA, in their LOR, and in their English usage. They were rather similar in the number of years that they spent studying English. On the basis of the information collected, we expected the Spanish group to be more proficient than the Mandarin group. On average, the Spanish group started learning English at a younger age, had spent more time in an English-speaking environment, and used English more extensively daily, compared to the Mandarin group.

Table 1. Summary of participants' background information

Language	Age at testing (years)	AOA (years)	LOR (months)	Years of study	Speaking (hr)	Listening (hr)	Reading (hr)
English	43 (26–55)	–	–	–	–	–	–
Spanish	42 (28–58)	8 (5–16)	113 (36–324)	9 (4–14)	27 (5–56)	40 (10–60)	27 (10–55)
Mandarin	22 (19–28)	16 (10–20)	28 (2–96)	8 (6–10)	10 (3–21)	32 (20–42)	19 (10–42)

Note. Data displayed are means, with ranges in parentheses; AOA = age of onset of acquisition; LOR = length of residency; speaking, listening, and reading = self-reported number of hours per week of English use.

Tasks and Stimuli

The same 21-year-old, female, native English speaker read materials for all tasks twice with very clear diction in the soundproof room of the Language Media Center at the University of Pittsburgh. The best version of the recording was selected by two phoneticians.

The +AM tasks were based on our own narrated version of the children's story *Frog Where Are You?*, by M. Meyer (Meyer, 1969). It consisted of 24 PowerPoint (version 14.3.8) slides, each containing an illustration and three to five sentences explaining it. Participants could listen only once to each slide.

The production +AM task consisted of 15 questions to elicit responses in relation to the story. The responses required sentence focus to be placed on the subject, verb, or object (five of each), plus nine distractors. Sample items are exemplified in (1):

(1) a. Focus on the subject: *Who is the dog? Is Bobby the dog?*
Expected answer: *No, TOBY is the dog.*
b. Focus on the verb: *What did the groundhog do to Bobby? Did it bite Bobby?*
Expected answer: *No, it SCARED Bobby.*
c. Focus on the object: *Who did the owl chase? Did the owl chase Toby?*
Expected answer: *No, the owl chased BOBBY.*
d. Distractor: *Did Toby fall out of the window? What happened to him?*
Expected answer: *Toby fell out of the window and the glass broke.*

The +AM perception task consisted of the same number of target questions and distractors as the +AM production task. It involved a forced-choice task in which participants had to select one of three possible answers that differed in the intonation contour, as illustrated in (2). Only one of the three intonations constituted a congruent answer in relation to the story; thus, the task required participants to match the meaning asked by the question with the F0 contour containing the appropriate prominent word.

(2) a. Focus on the subject: Question: *Did Bobby fall out of the tree?*
Answers: (a) *TOBY fell out of the tree*; (b) *Toby FELL OUT of the tree*; (c) *Toby fell out of the tree.*
b. Focus on the verb: Question: *That night, did Froggy sleep in his jar?*
Answers: (a) *That night, FROGGY escaped from his jar*; (b) *That night, Froggy ESCAPED from his jar*; (c) *That night, Froggy escaped from his jar.*
c. Focus on the object: Question: *Did the antlers belong to a moose?*
Answers: (a) *The antlers belonged to a DEER*; (b) *THE ANTLERS belonged to a deer*; (c) *The antlers belonged to a deer.*

The –AM production task involved the imitation of 15 decontextualized single utterances plus nine distractors, all unrelated in meaning,

as shown in (3). Notice that, unlike in the +AM production task, in the –AM production task, subjects were not required to produce a sentence whose prominence expressed a clear meaning. They could imitate the prominent word without understanding its contrastive meaning.

(3) a. Focus on the subject: *BOBBY cooked dinner.*
b. Focus on the verb: *Dad CAUGHT the balloon.*
c. Focus on the object: *Bobby threw the FOOTBALL.*

The –AM perception task had the structure of a forced-choice identification task, in which participants heard 15 unrelated and out-of-the-blue utterances differing in the focus placement plus nine distractors. Then they were asked to choose as the best match one of three possible sentences that differed only in intonation. Thus, this task required participants to match one of three F0 contours to the F0 contour of one isolated English sentence that focused speakers' attention more on the F0 contour form than on the meaning that the sentence expressed. The segmental content of these sentences was low-pass filtered at 450 Hz with 100 Hz of smoothing, using the pass Hann band from Praat (Version 5.1.43; Boersma & Weenink 2010). These values were selected after a pilot study in which two trained phoneticians listened to the low-pass filtered stimuli at 350 Hz, 450 Hz, and 550 Hz and were asked to transcribe the words and imitate the sentence intonation. The trained phoneticians could not transcribe the words reliably at 450 Hz but could easily repeat the intonation.

Procedure

To administer the two experiments, participants were alternatively assigned to start with either the –AM or the +AM tasks. In the –AM tasks, both production and perception stimuli were randomized for each participant. Before completing the –AM tasks, participants also received some training (a couple of slides showing the structure of the task) to become familiar with the identification of low-pass filtered utterances. The administration of the +AM tasks was preceded in all cases by listening to the story. After that, participants were alternatively assigned to complete either the perception or the production component first. The background questionnaire was administered between the two sets of tasks. Before the testing started, participants had to read a short text (that was also recorded) that explained the characteristics of the task.

Data Analysis

Measurements included the number of correct answers (perception tasks) and the relative duration, intensity, and mean F0 of vowels in target words (production tasks). Pitch was measured as the F0 maxima/F0 minima ratio on the stressed syllables of the subject, verb, and object in each sentence. For duration (and intensity), two ratios were computed. The duration (or intensity) of the stressed syllable of the subject was divided by that of the verb (i.e., S/V ratio), and the duration (or intensity) of the stressed syllable in the verb was divided by that of the object (i.e., V/O ratio). Computing these ratios eliminates two potential problems—namely, subjects' idiosyncratic differences, such as voices with different F0 and different speech rates, as well as the differences between scales in the measurements of pitch, duration, and intensity. Repeated-measures ANOVAs comparing tasks (+AM and –AM) and speakers' performance according to their L1 followed by post hoc tests were computed in SPSS.

RESULTS

Because *t* tests comparing possible order effects between perception and production tasks gave nonsignificant results, we begin by comparing the perception results in the +AM and –AM tasks. The –AM and +AM production results are presented in two separate sections, given the different nature of the collected responses. Although we had a comparable number and type of answers in the –AM task, in the +AM task, participants' responses varied greatly, and thus a more qualitative analysis ensues.

Perception of Focus in English: Effects of Access to Meaning, Focus Position, and Subjects' L1

To examine the effect of access to meaning and focus position on the perception of focus by L1 and L2 speakers, a repeated-measures ANOVA was performed on the number of correct answers obtained by each participant with task (+AM and –AM) and focus (subject, verb, and object) as the within-subjects factors and L1 (English, Mandarin, and Spanish) as the between-subjects factor. A summary of the ANOVA's results is presented in Table 2.

Results indicate that task played a significant role, showing that participants' performance was different between tasks, whereas the nonsignificant interaction of task and L1 indicate that this difference between tasks

Table 2. Subjects' perception of English focus: Results of a repeated-measures ANOVA with task (+AM tasks and –AM tasks) and focus (subject, verb, or object) as the within-subjects factors and subjects' L1 (English, Mandarin, or Spanish) as the between-subjects factor

Factor	Type III sum of squares	*df* num.	*df* den.	*F*	*p*
Task	7.934	1	37	4.750	.036
Task × L1	.798	2	37	0.239	.789
Focus	1.331	2	74	1.795	.173
Focus × L1	2.874	4	74	1.938	.113
Task × Focus	4.768	2	74	5.726	.005
Task × Focus × L1	6.367	4	74	3.824	.007

Note. Num. = numerator; den. = denominator.

was consistent across languages. A closer examination of the data shows that, in fact, participants obtained a higher number of correct responses in the –AM task than in the +AM task (–AM and +AM task means, respectively, for Mandarin: 4.69 [*SD* = .22] and 4.28 [*SD* = .24], English: 4.89 [*SD* = .22] and 4.69 [*SD* = .24], and Spanish: 4.38 [*SD* = .21] and 3.90 [*SD* = .23]). However, mean differences between tasks are so small (English: *M* = .20; Mandarin: *M* = .41; Spanish: *M* = .40) that, although they are statistically significant, they do not provide strong support for a clear difference between +AM and –AM tasks.

The nonsignificant results obtained for focus and for the interaction of focus and L1 (Table 2) indicate that, across tasks, there was not a specific focus position (subject, verb, or object) that participants of the different L1s perceived better or worse. However, within each task (–AM and +AM), some sentence positions were more difficult to perceive than others, as suggested by the significant interaction of task and focus. Additionally, the significant three-way interaction of task, focus, and L1 showed that the specific position differed across languages (see Figure 2). Tukey's honest significant difference (HSD) post hoc tests comparing focus and L1 within each task revealed that there were no significant differences between English and Mandarin speakers in any task nor in any sentence position, suggesting that L1 Mandarin speakers' perception of English focus did not differ significantly from that of the controls. In contrast, Spanish speakers did differ significantly from native English speakers in subject and verb positions but not in the object position (see Figure 2). In the –AM task (Figure 2, right), Spanish speakers significantly differed from the other two groups when the focus was on the subject (Spanish-English: $p = .032$; Spanish-Mandarin: $p = .051$). In the +AM tasks (Figure 2, left), L1 Spanish participants significantly differed from English native speakers when the focus was on the verb (Spanish-English: $p = .034$), and a marginally significant p value of .066

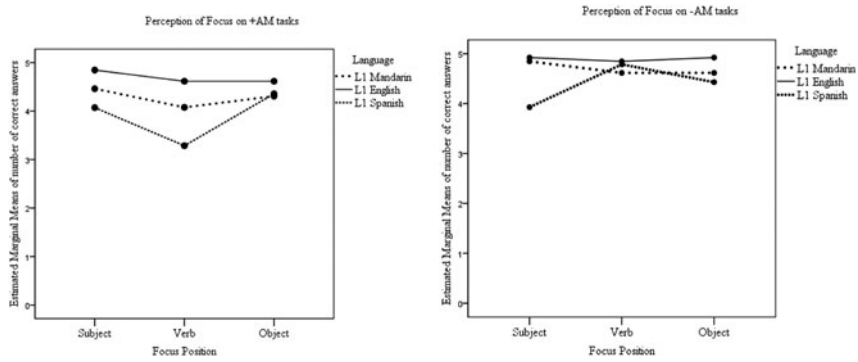

Figure 2. Subjects' perception of focus organized by participants' L1 and focus position. Left: results for the +AM tasks. Right: results for the –AM tasks.

also suggests that the perception of focalized subjects proved more difficult for Spanish speakers than for controls.

To assess the overall effect of the speakers' L1 across tasks and focus positions, Tukey's HSD multiple comparisons tests were performed. Results confirmed that L1 Mandarin speakers' perception of English focus did not significantly differ from that of English native speakers, $p = .457$. However, Spanish speakers' perception did, as they showed a significantly higher number of errors than English native speakers, $p = .034$.

In summary, the significant effects of task were too small to give reasonably strong evidence in support of our prediction that facilitating access to meaning would increase L1 transfer. However, significant differences among language groups were consistent and supported our predictions regarding L1 transfer. Overall, Mandarin speakers did not differ significantly from English controls, whereas Spanish speakers did. Moreover, Spanish speakers were more prone to err at perceiving focus in subject and verb positions than in object position.

Production of Focus: Sentence Imitation Task (–AM Task)

In the imitation (-AM) task, speakers behaved similarly in their use of pitch and intensity regardless of their L1. As for pitch (see Figure 3), speakers of each language showed a large pitch range on the subject and a reduced pitch range on the verb when the sentence that they were imitating had a focalized subject (Figure 3, left). As a result, the pitch range of the subject was expanded and stood out in comparison to that of the adjacent verb. When the focus was on the verb (Figure 3, center), speakers increased the pitch range of the stressed syllable of the focalized

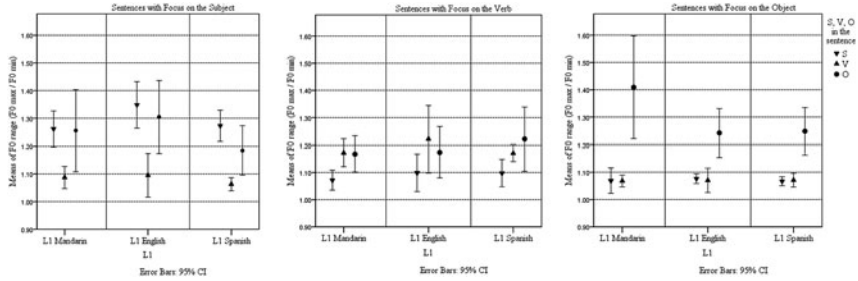

Figure 3. F0 normalized values obtained for each language group in the sentence imitation task (–AM tasks). Results are displayed by focus position. Left: focus on the subject. Center: focus on the verb. Right: focus on the object. S = subject, V = verb, O = object.

word in comparison to that of the subject. Finally, when the object was focalized, the pitch range of the accented syllable was expanded in comparison to the pitch range on other constituents (Figure 3, right). Similar patterns were observed for intensity, such that the stressed syllable that was focalized had a larger relative intensity than the stressed syllables of nonfocalized constituents.

In general, speakers from the three language groups enlarged the duration of the word in focus in comparison to the words in postfocal position, maintaining in this way a duration difference between focus and background (see Figure 4). There were, however, some subtle differences. When the subject was in focus position, the duration of the subject was not as different from that of the verb in Spanish when compared to controls (Figure 4, left). In comparison to controls, Mandarin speakers reduced the difference in duration between the verb and the focalized object by stretching the stressed syllable of the verb (Figure 4, right). To assess the statistical significance of these patterns, repeated-measures ANOVAs with subjects' L1 (between-subjects factor) and focus

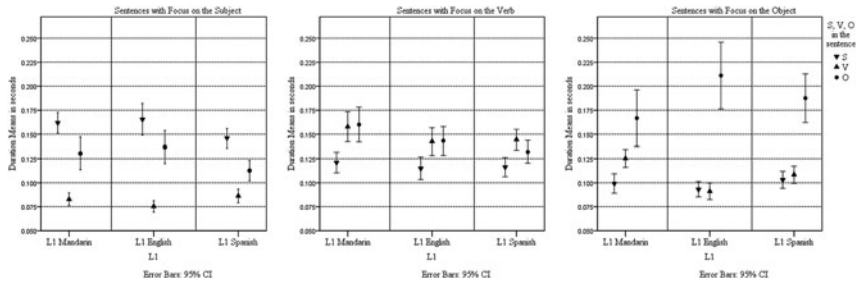

Figure 4. Mean duration values in the sentence imitation task (–AM tasks). Results are displayed by focus position. Left: focus on the subject. Center: focus on the verb. Right: focus on the object.

position (repeated measure), were run separately for each focus position (subject, verb, and object) on the pitch, duration, and intensity ratios.

The nonsignificant results obtained for pitch and intensity confirm that all speakers, independently of their L1, were successful at imitating the use of these parameters to mark focus; namely, L2 speakers managed to enlarge the pitch range or the intensity differences in the stressed syllables of the focalized words and to reduce those same ranges in the stressed syllables of nonfocalized words. However, significant crosslanguage differences were obtained for duration in both sentences with the focus on the subject and those with the focus on the object (subject: S/V ratio: $F(2, 30) = 5.815$, $p = .010$, V/O ratio: $F(2, 30) = 4.547$, $p = .024$; object: S/V ratio: $F(2, 30) = 6.579$, $p = .004$, V/O ratio: $F(2, 30) = 12.102$, $p < .001$). Tukey's HSD multiple comparisons test revealed that Spanish speakers performed differently from natives when the subject was focalized (S/V ratios: Spanish-English = .012, Mandarin-English = .858), whereas L1 Mandarin speakers performed differently from native speakers when the focus was on the object (V/O ratios: Mandarin-English = .000, Spanish-English = .865). These differences, however, did not result in overlapping distributions, which would cancel out the durational contrast between focus and background; rather, they resulted in only a reduction of the contrast.

In summary, L2 speakers were able to imitate the contrast between focus and background. This contrast was conveyed by enlarging the pitch range, duration, and intensity of the stressed syllable of the focalized word and by reducing the same parameters in the stressed syllables of the adjacent words (background). The minor crosslanguage differences in the use of duration did not affect the expression of this contrast per se but resulted in a subtle reduction of the degree of contrast for only one dimension (i.e., duration).

Production of Focus: Answering Questions in Relation to a Story (+AM Task)

Given that this was an open-ended task, which was also, to some extent, dependent on participants remembering aspects of the story, a greater degree of interspeaker variability vis-à-vis the imitation task was expected. The first apparent difference between controls and L2 speakers was that native English speakers tended to answer in complete sentences. For example, to the question *Who is the dog? Is Bobby the dog?*, English speakers' most frequent answer was *No, Toby is the dog*, with the expected focus on the subject *Toby*. In contrast, some Spanish speakers offered incomplete sentences or single words, such as *No, Toby*. Following the criterion that a complete sentence includes a subject, a verb, and

an object, participants' answers were classified by a native English speaker as either complete or incomplete. Then the number of complete sentences was divided by the number of questions answered. For example, participant S1 answered with five complete sentences to the five questions designed to elicit focus on the subject, obtaining a ratio of 1. Mean ratios, which are displayed in Table 3, indicate that English speakers consistently produced complete sentences. First language Mandarin speakers displayed a strong tendency to answer with complete sentences, whereas Spanish speakers did so less often. A repeated-measures ANOVA on the aforementioned ratios with focus position (subject, verb, and object) as the repeated measure and participants' L1 (Mandarin, English, and Spanish) as the between-subjects factor showed that the differences in the proportion of complete answers did not reach significance.

Additionally, groups differed in their degree of fluency. As expected, native speakers tended to pause less often than L2 speakers. To compare the use of pauses among the three groups, we classified the pauses into two types—namely, those pauses placed after the focalized word (e.g., *TOBY* [pause] *is the dog*) versus those found elsewhere (*No, the-* [pause], *they went TO THE WOODS*). Means calculated from the ratios of the two pause types (see Table 4) showed that English speakers paused less often than L2 speakers and that Spanish speakers paused after focus more often than any other language group. Repeated-measures ANOVAs on the ratios showed that speakers' differences on pauses elsewhere did not reach significance, $F(2, 28) = 0.942$, $p = .402$, but differences in the use of postfocal pauses did, $F(2, 28) = 7.627$, $p = .002$. Post hoc tests confirmed that Spanish speakers used significantly more pauses after focus than any other language group.

The last step of the analysis involved the acoustic analysis of the complete answers produced by each speaker. Those utterances were manually labeled according to the following criteria: First, in sentences with focus on the subject and on the verb, the stressed vowel was marked as "F," and the stressed vowel of the first full word after focus as "Post-F." Second, in sentences with focus on object, the label "Pre-F" was placed on the stressed vowel of the word preceding the focalized word.

Table 3. Mean ratios of complete vs. incomplete answers by focus position and L1 (+AM tasks)

Focus	English	Mandarin	Spanish
Subject	1.0 (0.0)	1.0 (0.0)	.908 (.175)
Verb	1.0 (0.0)	.960 (.084)	.969 (.075)
Object	1.0 (0.0)	1.0 (0.0)	.831 (.335)

Note. Standard deviations are given in parentheses.

Table 4. Mean ratio of postfocal (Post-F) vs. other pauses displayed by L1 and focus position

	English		Mandarin		Spanish	
Focus	Post-F	Elsewhere	Post-F	Elsewhere	Post-F	Elsewhere
Subject	.025 (.07)	.02 (.07)	.06 (.096)	.24 (.18)	.22 (.21)	.17 (.16)
Verb	.025 (.07)	.20 (.28)	.06 (.96)	.28 (.21)	.15 (.17)	.33 (.17)
Object	n/a	.08 (.14)	n/a	.16 (.12)	n/a	.17 (.20)

Note. Standard deviations are given in parentheses.

Then the relative duration, intensity, and F0 ratios were calculated in the same way as in the production portion of the –AM task. Given that, in the +AM task, the number of sentences per subject and per language group was different (we remind the reader that not all participants answered in full sentences), the nonparametric Kruskal-Wallis test was used to examine the effect of the speakers' L1 in their production of pitch, duration, and intensity ratios. Results showed nonsignificant differences for duration and intensity (for duration: on subject, $p = .065$, on verb, $p = .487$, and on object, $p = .111$; for intensity: on subject, $p = .658$, on verb, $p = .114$, and on object, $p = .088$) and significant crosslanguage differences for pitch ratios in sentences with focus on the subject, $p = .007$, and with focus on the object, $p = .004$. Post hoc tests revealed that Mandarin and Spanish speakers significantly differed from each other but not from controls (pitch on subject: English-Mandarin, $p = .248$, English-Spanish, $p = .112$, and Spanish-Mandarin, $p = .001$; pitch on object: English-Mandarin, $p = .840$, English-Spanish, $p = .113$, and Spanish-Mandarin, $p = .038$). Figure 5 illustrates where the differences are. It can be seen (Figure 5, left) that

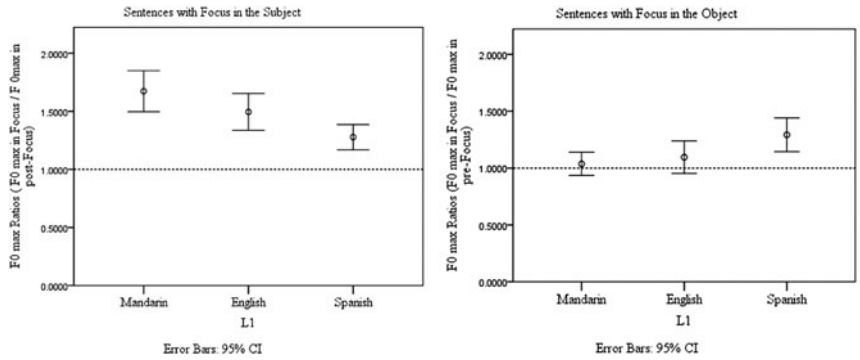

Figure 5. F0 ratios obtained for each language group in the contextualized task (+AM tasks). Results are displayed by focus position. Left: focus on the subject. Right: focus on the object.

Spanish speakers reduced the contrast between focus and background in sentences with focus on the subject, whereas they amplified this contrast when the focus was on the object (Figure 5, right).

In summary, in comparison to English controls, L2 speakers answered less often in complete sentences and produced more "elsewhere" pauses. These differences, although they did not reach significance, showed that, in general, English speakers' answers were more fluent than those of L2 speakers. Differences among language groups appeared in participants' use of pauses that were placed after focus. Whereas Mandarin speakers did not differ from English controls, Spanish speakers placed more pauses after a word in focus than any other language group, which supports our L1 transfer hypothesis. As for participants' use of the acoustic dimensions of duration, intensity, and pitch, only differences in the use of pitch were found. In particular, Spanish speakers overamplified the pitch range of the focus on object position, where L1 transfer predicted a good performance, and reduced it in subject position, where L1 transfer predicted more errors.

DISCUSSION

Overall, L2 speakers did comprehend and use the acoustic cues of pitch, duration, and intensity to produce corrective focus. Nevertheless, differences found between tasks and among language groups allow us to examine the predictions regarding L1 transfer, access to meaning, and production versus perception. As for L1 transfer, we predicted positive L1 transfer for Mandarin speakers and negative L1 transfer for Spanish speakers. Our results consistently support these predictions. In the perception tasks, Spanish speakers significantly differed from controls, as opposed to Mandarin speakers. Moreover, in the +AM production task, only Spanish speakers differed significantly from controls in the insertion of pauses after the focalized word. Because focus at the end of the sentence, which is the strategy used in Spanish, is usually followed by a pause, this pause can be interpreted as a redundant cue to the focalized word in Spanish. Therefore, a pause insertion after the focalized word in L2 English can be interpreted as evidence for L1 influence.

These differences across groups could also be attributed to proficiency levels. As the reader may recall, the two groups of learners tested here, albeit highly advanced, differed in their experience with the target language. Direct comparisons between L2 groups that differ in terms of their L1 are not always straightforward, and our L1 Spanish speakers had more experience with the target language, at least if we consider the variables taken into account in the present study. In spite of that, it was the Mandarin group that matched controls more closely. Thus, we interpret this as evidence that language-specific strategies to express

contrastive focus (and not proficiency) are behind the differences obtained in the results.

Beyond the group differences, our L1 transfer predictions referred to the types of errors performed by Spanish speakers. In Spanish, focus is placed at the end of the sentence; thus Spanish speakers were expected to be biased toward perceiving and producing focus in that position. The results provide strong evidence in support of this prediction. In perception, Spanish speakers differed from controls when the focus was on the subject and on the verb. In production, they reduced the highlight-background contrast when the word in focus was in subject position by shrinking its pitch range (+AM task) and duration (–AM task) in comparison to native English speakers' productions. However, when the object was focalized, Spanish speakers tended to overemphasize the highlight-background contrast by enlarging the pitch range of the object more than English speakers did (+AM task). Thus, these results show that Spanish speakers differed more from controls in perception and production when the focus was on the subject, whereas they resembled English participants more when the focalized word was the object, providing support for our predictions on error types.

Regarding access to meaning, we predicted that making meaning more or less accessible by manipulating context would have an effect on L1 transfer: Access to meaning was expected to increase L1 transfer. This effect would translate into a difference in performance between +AM and –AM tasks, which was expected to be particularly clear in Spanish. Results showed that manipulating context indeed had an effect on L1 transfer. Whereas, in perception, participants' better performance in –AM than in +AM tasks cannot be interpreted as strong evidence supporting our hypothesis (albeit significant, results are based on very small differences), more interesting conclusions may be drawn from production. Although, in general, L2 speakers were able to use the relevant acoustic cues to highlight the word in focus in both +AM and –AM production tasks, they consistently produced sentences in the +AM task with pauses that did not appear in the –AM task. Of particular interest are the pauses that Spanish speakers inserted after the focalized subject or verb, which constitutes evidence of L1 transfer. As a consequence, production in the +AM task sounded less fluent vis-à-vis the –AM task, in spite of the fact that the sentences in both tasks were of similar length and syntactic complexity. Moreover, in the +AM task, subjects looked at the illustration that portrayed the action to be described in the answer, which eased the memory load required to retell the story. Thus, the overall performance by Spanish speakers in the +AM production task suggests that facilitating access to meaning does increase L1 transfer.

Our final prediction concerned production versus perception. We predicted that if perception led production, L1 transfer would be greater

in production and that any strategy that affected L1 transfer in perception would affect production as well. Results suggest that L1 transfer was more evident in production than in perception, especially for Spanish speakers, which supports the claim that perception leads production. Moreover, L2 speakers performed better in –AM than in +AM tasks. Because this was a clear tendency in production, and because it was also supported by a small, albeit statistically significant, change in perception, there is no evidence against the prediction that facilitating access to meaning increases L1 transfer in both perception and production.

Beyond testing specific predictions, our study intended to contribute to the larger field of L2 speech by exploring how facilitating access to meaning by manipulating context affected L1 transfer, and it showed that L1 transfer increased with access to meaning, especially in production. As our results revealed, there are interesting parallels between our findings and previous literature. In particular, L2 learners performed worse in tasks in which meaning was made more accessible by contextualizing sentences within a story in comparison to tasks in which meaning was less accessible and attention to acoustic cues could be sufficient to perform the task. This is consistent with studies on L2 segment and stress perception that have shown that language experience affects higher levels of processing such as speech-specific selective attention. The present study contributes as well to research on L2 intonation by showing that not only do L1 transfer effects disappear when listening to nonspeech stimuli, as shown by Grabe et al. (2003), but they are modulated as well by different levels of access to meaning, which indicates that increasing the levels of processing leads to increased L1 effects. Finally, our results also suggest that, to fully understand the acquisition of L2 prosody, tasks aimed at different processing levels should be taken into consideration.

CONCLUSIONS

The two tasks and four experiments discussed here were designed to test the hypothesis that acquiring intonation in a L2 not only is an issue of learning to perceive and produce the target melody but, crucially, involves a new mapping between form and meaning that is affected by L1 transfer. This hypothesis contends that, in turn, L1 transfer can be molded by manipulating access to meaning through context. Our results showed that L2 speakers performed worse in the story-related tasks, in which changes in the F0 contour had implications for the correct understanding of the story, than in the –AM tasks, in which understanding the meaning of the sentence was not crucial for either the perception or the correct imitation of the sentence. These results,

which were particularly strong for the L1 Spanish group and in production tasks, revealed that facilitating access to meaning increased L1 transfer.

REFERENCES

Best, C. (1994). The emergence of native language phonological influence in infants: A perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: Transition from speech sounds to spoken words* (pp. 167–264). Cambridge, UK: Cambridge University Press.

Best, C., & Tyler, M. (2007). Nonnative and second language speech perception: Commonalities and complementarities. In O. Bohn & M. Munro (Eds.), *Language experience in second language learners* (pp. 13–34). Amsterdam, the Netherlands: Benjamins.

Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of Mandarin tone 2/tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics*, *18*, 37–49.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 5.1.43) [Computer software]. Retrieved from http://www.praat.org/

Bolinger, D. (1989). *Intonation and its uses*. Stanford, CA: Stanford University Press.

Chen, A., Rietveld, T. C. M., & Gussenhoven, C. (2001). Language-specific effects of pitch range on the perception of universal intonational meaning. In P. Dalsgaard, B. Lindberg, H. Benner, & Z.-H. Tan (Eds.), *EUROSPEECH-2001* (pp. 1403–1406). Retrieved from http://www.isca-speech.org/archive/eurospeech_2001/e01_1403.html

Cruttenden, A. (1997). *Intonation*. Cambridge, UK: Cambridge University Press.

Cruz-Ferreira, M. (1984). Perception and interpretation of non-native intonation patterns. In M. P. R. van der Broecke & A. Cohen (Eds.), *Proceedings of the 10th International Congress of Phonetic Sciences* (pp. 565–569). Dordrecht, the Netherlands: Foris.

Cruz-Ferreira, M. (1989). A test for non-native comprehension of intonation in English. *International Review of Applied Linguistics in Language Teaching*, *27*, 23–39.

Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language*, *36*, 406–421.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness." *Journal of the Acoustical Society of America*, *110*, 1606–1618.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress "deafness" in simultaneous French-Spanish bilinguals. *Cognition*, *114*, 266–275.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress "deafness": The case of French learners of Spanish. *Cognition*, *106*, 682–706.

Face, T. (2001). Focus on peak alignment in Spanish intonation. *Probus*, *13*, 233–246.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J., & Frieda, E. (1997). Amount of native language use (L1) affects the pronunciation of an L2. *Journal of Phonetics*, *25*, 169–186.

Grabe, E., Rosner, B. S., García-Albea, J. E., & Zhou, X. (2003). Perception of English intonation by English, Spanish, and Chinese listeners. *Language and Speech*, *46*, 375–401.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge, UK: Cambridge University Press.

Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones*. Cambridge, UK: Cambridge University Press.

Ladd, R. (2008). *Intonational phonology*. Cambridge, UK: Cambridge University Press.

Liu, F. (2009). *Intonation systems of Mandarin and English: A functional approach* (Unpublished doctoral dissertation). University of Chicago.

McGory, J. T. (1997). *Acquisition of intonational prominence in English by Seoul Korean and Mandarin Chinese speakers* (Unpublished doctoral dissertation). The Ohio State University, Columbus.

Meyer, M. (1969). *Frog, where are you?* New York, NY: Dial Press.

Nava, E., & Zubizarreta, E. (2009). Order of L2 acquisition of prosodic prominence patterns: Evidence from L1 Spanish/L2 English speech. In J. Crawford, K. Otaki, &

M. Takahashi (Eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)* (pp. 175–187). Somerville, MA: Cascadilla Press.

Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2003). Linguistic relevance of duration within native language determines the accuracy of speech-sound duration processing. *Cognitive Brain Research*, *16*, 492–495.

Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain and Language*, *92*, 26–32.

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.

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994). Perceptual learning of non-native speech contrasts: Implications for theories of speech perception. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception* (pp. 121–166). Cambridge, MA: MIT Press.

Toledo, G. (1989). Señales prosódicas de foco [Prosodic focus marking]. *Revista Argentina de lingüística*, *5*, 205–230.

Vallduví, E. (1991). The role of plasticity in the association of focus and prominence. In Y. No & M. Libucha (Eds.), *ESCOL '90: Proceedings of the Seventh Eastern State Conference in Linguistics* (pp. 295–306). Columbus, OH: The Ohio State University.

Wang, B., & Xu, Y. (2011). Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of Phonetics*, *39*, 595–611.

Watson, D. G., Arnold, J. E., & Tanenhaus, M. K. (2008). Tic tac toe: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, *106*, 1548–1557.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7*, 49–63.

Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *25*, 121–133.

Xu, L. (2004). Manifestation of informational focus. *Lingua*, *114*, 277–299.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, *27*, 55–105.

Xu, Y. (2001). Fundamental frequency peak delay in Mandarin. *Phonetica*, *58*, 26–52.

Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.

Zubizarreta, M. L. (1998). *Prosody, focus, and word order*. Cambridge, MA: MIT Press.

Zubizarreta, M. L., & Nava, E. (2011). Encoding discourse-based meaning: Prosody vs. syntax. Implications for second language acquisition. *Lingua*, *121*, 652–669.