DISCOURSE INTONATION IN L2 ACADEMIC PRESENTATIONS: A PILOT STUDY

Larissa Buss, Concordia University

Walcir Cardoso, Concordia University

Sara Kennedy, Concordia University

Intonation is an important cue to the organization of oral discourse, being used by first language (L1) English speakers to signal topic shifts and relationships among parts of speech (Wennerstrom, 1994, 1998). This pilot study investigated how four second language (L2) graduate students' use of intonation as an organizational device developed naturalistically during their first six months immersed in an L2 environment. The participants were recorded delivering four short presentations at approximately two-month intervals. The recordings were analyzed for two features: *paratones* (extra high pitch at the beginning of a new discourse topic) and *mid-utterance pitch boundaries* (which should indicate the connection between phrases with non-low pitch). Two presentations given by native English speakers were also analyzed for comparison. Overall, the L2 participants adequately produced non-low pitch boundaries within utterances, but their F0 peaks at topic shifts were considerably less prominent than those observed in the native-speaker data. Three participants' use of intonation changed over time, either improving on the features analyzed or displaying a U-shaped pattern of development.

INTRODUCTION

Intonation can have a variety of communicative functions, which include conveying the illocutionary force of an utterance, expressing attitudes and emotions, and building rapport (Chun, 2002; Pickering, 2001). No less important is its role in organizing discourse (Wennerstrom, 1998). In English, variations in pitch are used alongside syntactic and lexical devices to create a hierarchical structure in spoken discourse, highlighting important information and signaling relationships among constituents. However, given that intonation patterns vary across languages, it may take time, exposure, and practice for learners to master the subtleties of discourse intonation in their second language (L2). This pilot study analyzed changes in L2 graduate students' use of intonation as an organizational device, specifically paratones and utterance-medial pitch boundaries, over two academic terms.

Description of the Intonation Variables

Intonation is an important cue to the organization of large stretches of speech. A new discourse topic usually begins with a high peak on the first prominent syllable and a high pitch overall, followed by "a series of lesser peaks" (Yule, 1980, p. 36). When the topic comes to an end, there is a lowering of the pitch level and often a slowing down, followed by a lengthy pause (Yule, 1980). These cues segment discourse into macro-units that are analogous to paragraphs in

writing. Brown (1977) referred to them as *paratones*, but they have also been called *phonological paragraphs*, *pitch sequences*, *sequence chains*, and *intonational paragraphs*.

Besides marking the macro-structure of discourse, intonation is also used to indicate the relationship between parts of speech. A low pitch boundary, observed at the end of declarative statements, is generally associated with finality or closure (Wennerstrom, 2001). It indicates that the speaker has finished what they wanted to say. In an interaction, a marked fall in pitch (low) signals the speaker's intention to give away the turn. On the other hand, high-rising, low-rising, partially falling, and plateau boundaries (all referred to in this study as "non-low") indicate interdependency between the phrases in which they occur and those that follow. That is, they anticipate subsequent phrases for their full and correct interpretation (Wennerstrom, 2001).

Research on the Intonation Variables

Researchers have reported similar findings on how native speakers use paratones and pitch boundaries and on the phonetic/phonological cues associated with these features. Swerts and Geluykens (1994) analyzed the features in the context of spontaneous monologues in Dutch. It was found that the fundamental frequency (F_0) peaks of topic-introducing clauses were higher than the F_0 peaks of the following clauses. Moreover, there was a significant association of low-ending contours with topic finality and of high-ending contours with non-finality. Thompson (2003) described the use of paratones to signal the larger-scale organization of authentic and pedagogically prepared academic lectures in English. The paratones were characterized by extra high pitch on the first prominent syllable and extra low pitch on the last, sometimes ending with decreased volume or speed, laryngealization, and a long pause.

Grosz and Hirschberg (1992) found evidence that listeners pick up on these intonational cues. Seven listeners were asked to label the discourse structure of a news story, either from text alone (with most punctuation removed) or from text and speech. For both modalities, phrases labeled as initiating discourse segments were produced with a larger pitch range than other utterance-initial phrases. In another study (Lehiste, 1975), a group of 30 listeners were asked to determine the context of production of randomized recorded sentences (if isolated, paragraph-initial, medial or final). The results revealed that sentences produced with high F0 peaks were consistently identified as being paragraph initial. This was interpreted as evidence that listeners have certain expectations regarding the intonation of sentences within an orthographic paragraph, which is a macro-unit comparable to the paratone.

In L2 speech, listener expectations related to discourse intonation are not always met. Tyler, Jefferies, and Davies (1988) observed that the intonation contours of Korean and Chinese teaching assistants (TAs) sometimes ended in a falling pitch, even within sentences, which obscured markings of continuation and finality. According to the authors, this also made the speakers sound abrupt, cold, and impatient. Likewise, Wennerstrom (1994) found that Japanese and Thai subjects used low pitch boundaries in the middle of utterances, failing to establish the link between semantically related phrases. Furthermore, they did not increase pitch range to signal new topics like L1 English and Spanish speakers did. Similar distinctions between native and nonnative speakers were reported in Pickering (2004). In this study, American TAs were found to consistently use the phonological cues of paratones to organize their discourse, whereas the same cues in the speech of Chinese TAs often did not correspond to semantic or structural boundaries in their discourse.

There is some evidence that inaccurate use of discourse intonation may undermine effective communication. Tyler and Davies (1990), for example, described an interaction between a Korean TA and an undergraduate student where communication failed due to problematic intonational cues. Furthermore, Wennerstrom (1998) found a significant relationship between paratone accuracy and the speaking scores of L1 Mandarin speakers lecturing in L2 English. The author also noticed that speakers with high scores tended to use non-low intonation in utterance-medial boundaries, as did native speakers, whereas the participant who received the lowest score produced many more misleading low boundary tones.

Despite the importance of paratones and pitch boundaries, little is known about how these features are acquired by L2 learners immersed in an English-speaking environment. Previous research suggests that learners can improve their pronunciation even without explicit instruction, especially in their first year of immersion (Derwing, Thomson, & Munro, 2006; Flege, 1988). However, it is not clear whether this applies to features of discourse intonation, and, if so, how these features develop over time in L2 speech. The present study contributes to our understanding of how discourse intonation develops naturalistically in L2 English speakers during their first six months immersed in an L2 context. The genre chosen for analysis was the academic presentation.

Research Questions

There are four research questions in this study, two related to paratones and two related to pitch boundaries:

1. Do L2 English/L1 Mandarin speakers differ from L1 English speakers in their use of paratones in oral presentations? If so, how?
2. Do L2 English/L1 Mandarin speakers differ from L1 English speakers in their use of utterance-medial pitch boundaries in oral presentations? If so, how?
3. Does the L2 speakers' use of paratones change across two semesters of immersion in an academic, English-speaking environment? If so, how?
4. Does the L2 speakers' use of utterance-medial pitch boundaries change across two semesters of immersion in an academic, English-speaking environment? If so, how?

METHODS

Participants

This study is part of a larger project examining longitudinal changes in the speaking skills of nonnative graduate students at a Canadian university. The participants here were four L1 speakers of Mandarin who were living abroad for the first time and who had arrived in Canada approximately one month before data collection. They did not receive any formal instruction on pronunciation or academic speaking during the 26-week study. More information is given in Table 1, with the participants' names changed to preserve confidentiality.

Table 1

Participant Information

Name	Gender	Age	Test Score	Program
Xu	F	22	IELTS 7.0	Graduate Diploma in Instructional Technology
Yang	F	25	TOEFL 92	Ph.D. in Civil Engineering
Wei	M	23	IELTS 6.5	M.A.Sc. in Mechanical Engineering
Chen	M	28	TOEFL 83	Master of Applied Computer Science

Two Canadian native English speaker graduate students were also included in the study as controls, a male aged 30 (here called Peter) and a female aged 39 (here called Kim). They were unaware of the purpose of the study.

Data Collection

Speech samples from the L2 graduate students were collected four times over a period of 26 weeks at approximately two-month intervals. A few days before each data collection session, the participants were asked to prepare a five-minute presentation explaining a key concept or term in their field to an imagined audience of first-year undergraduate students. The two native controls were recorded giving in-class presentations for a course. Each of them gave one presentation of approximately five minutes where they summarized and commented on part of a research paper that they had read.

Coding and Analysis

The presentations were transcribed and coded for intonational phrases, phonological utterances, and discourse topics. The first two constructs were conceptualized following Nespor and Vogel's (1986) model of prosodic phonology. According to this model, the intonational phrase (*I*) is the domain of an intonation contour, and its boundaries coincide with the places where pauses may be introduced in an utterance. An *I* is delimited by the boundaries of a root sentence (Emonds, 1976), which is roughly equivalent to a main clause or independent clause –one which is not embedded within a larger structure. Nonetheless, a root sentence does not necessarily form a single *I*. The division of a string into *I*s varies as a function of several factors, including length, rate of speech, emphasis, and style.

Besides a coherent intonation contour and pauses, other phonetic/phonological cues that helped to identify the boundaries of intonational phrases were: lengthening of the final syllable, acceleration at the beginning of a new phrase, a change in the pitch level or pitch direction of an unaccented syllable, and the presence of at least one pitch accent in the intonation phrases created (Cruttenden, 1986).

The phonological utterance (*U*) is the largest constituent in Nespor and Vogel's (1986) hierarchy, and it is formed by one or more intonational phrases. It usually corresponds to the string dominated by the highest node of a syntactic tree, or what is generally recognized as a "sentence". Although *U* can be defined using syntactic information, phonological and semantic factors may cause the prosodic constituent to be restructured, such that the end result may not necessarily be isomorphic to any syntactic constituent.

The pitch boundaries of *I*s situated in the middle of a *U* (i.e., in non-final position) were the aspects of interest in this study. These phrases were expected to end in non-low pitch boundaries, which express "non-finality" and dependency on subsequent phrases (Wennerstrom, 1998). Following Wennerstrom (1998), the pitch boundaries were classified as "low" or "non-low" based on auditory impression and observation of F_0 contours on Praat (Boersma & Weenink, 2014). The final F_0 value of the contour was also noted to determine whether the pitch dropped to the bottom of the speaker's pitch range (Wennerstrom, 2001). Finally, the percentage of non-low boundaries was calculated for each presentation. Truncated (i.e., uncompleted) intonational phrases and fragments that did not contain at least one pitch accent, such as fillers occurring in isolation (e.g., 'uh', 'eh', 'and', 'so'), were excluded from the analysis.

Discourse topics were identified mainly on the basis of semantic content, as "an aggregate of coherently related events, states, and referents" (Chafe, 1994, p. 121). Coding was also aided by the presence of "macro-markers" (Chaudron & Richards, 1986) such as "The first thing I'm going to look at is…" and "The second thing we need to do is…". When used, these markers made the discourse structure very explicit. Yet, most of the time, the speakers transitioned directly from one topic to another or signaled the shift using "micro" discourse markers such as "Also", "As well", and "Then" (Chaudron & Richards, 1986).

Based on the method used in Swerts and Geluykens (1994), the analysis of the paratone was done by measuring the first F_0 peak in each paratone-initial and paratone-final *I* on Praat (i.e., the first F_0 peak in the first and last *I*s of each paratone). Then, the value found for the final *I* of one paratone was subtracted from the value found for the initial *I* of the subsequent paratone. This resulted in measures of pitch increase at rhetorical junctures. Each participant's measures were converted into percentages and averaged.

RESULTS

Paratones

The analysis of F_0 peaks at the beginning of paratone-initial and paratone-final *I*s revealed considerably higher percentages of pitch increase for the L1 controls than for L2 speakers. As shown in Figures 2 and 3, the native speakers increased their pitch by 48% (Peter) and 57% (Kim) to signal topic shifts, whereas the nonnative speakers' increases ranged from 7% to 32%, with a total average of 19%. Large differences between native and nonnative speakers were also observed in Wennerstrom (1994, 1998). As in Wennerstrom (1998), the female control produced a higher increase than the male, but this gender difference was not found for the L2 participants. Figure 1 shows an example of pitch increase from the end of topic 7 to the beginning of topic 8 in Peter's presentation.
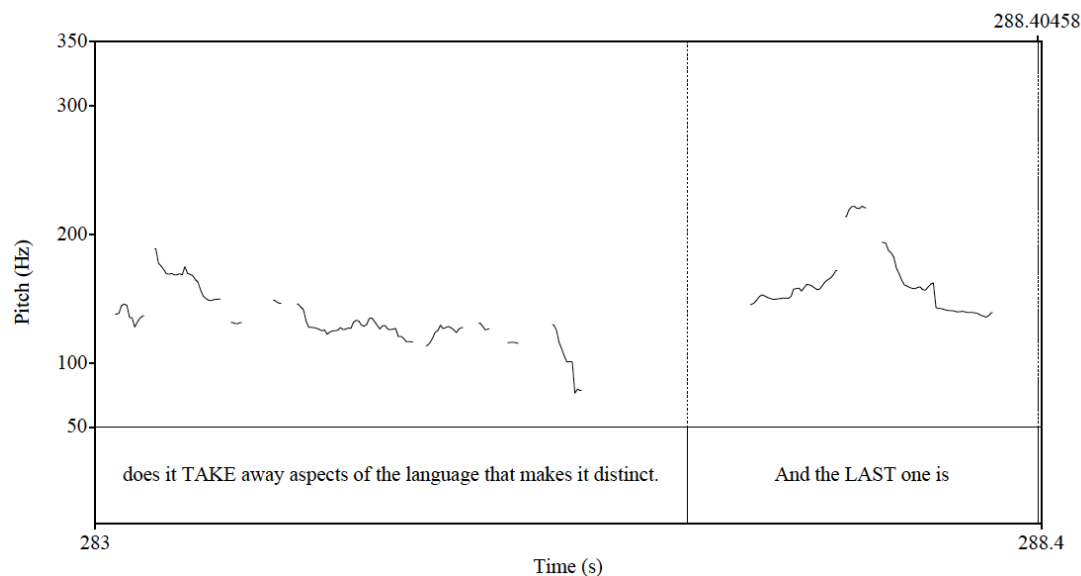
Figure 1. Pitch increase at a topic shift in Peter's presentation (L1 English).

Longitudinally, results seemed to differ for male and female L2 speakers. Xu and Yang, who had higher proficiency test scores than the males (see Table 1), started the study with low percentages of pitch increase at P1, but seemed to improve over time (see Figure 3). As displayed in Figure 4, Wei and Chen started at higher percentages of pitch increase than the females and closer to their L1 counterpart. However, Chen's means dropped at P2 and P3, then increased again at P4. Wei, on the other hand, did not exhibit any major changes.
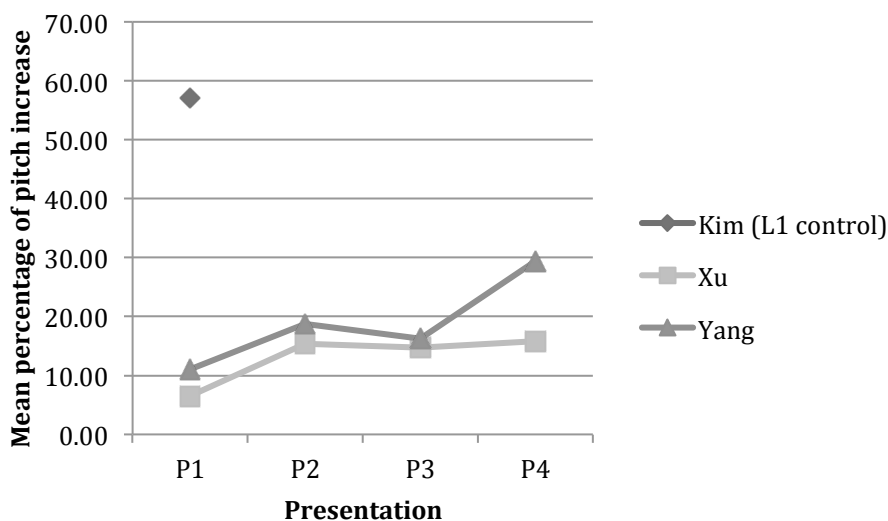

Figure 2. Mean percentage of pitch increase at topic shifts for female participants.
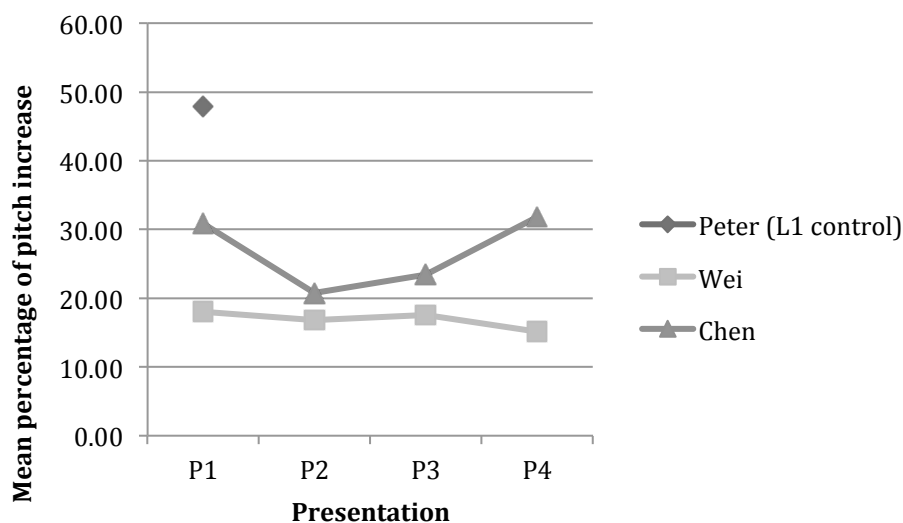
Figure 3. Mean percentage of pitch increase at topic shifts for male participants.

Utterance-Medial Pitch Boundaries

In the native speaker presentations, all mid-utterance boundaries were non-low, as displayed in Table 2. This feature was almost at ceiling level for L2 speakers as well, as most presentations exhibited over 90% of non-low boundaries in the context analyzed. The only exception was the first presentation by Chen, where 82% of the boundaries were non-low. Figure 4 shows the F0 contour for a problematic utterance produced by Chen in his first presentation (P1). The four pitch boundaries are low (↓), plateau (→), low (↓), and partially falling (↘). The two low boundaries misleadingly indicate closure rather than connection between the phrases.

Table 2

Percentages of Non-Low Utterance-Medial Boundaries

Participant	P1	P2	P3	P4
Peter (control)	100	-	-	-
Kim (control)	100	-	-	-
Xu	91	97	95	97
Yang	92	94	100	99
Wei	97	96	100	100
Chen	82	93	100	100

Over time, Chen improved his use of pitch boundaries, as his percentage of non-low boundaries increased to 93% at P2 and 100% at P3. It is possible that Xu and Yang also experienced slight improvement before reaching the ceiling. Overall, however, the accuracy of utterance-medial boundaries was high, therefore no major differences from the native controls or longitudinal changes were observed.
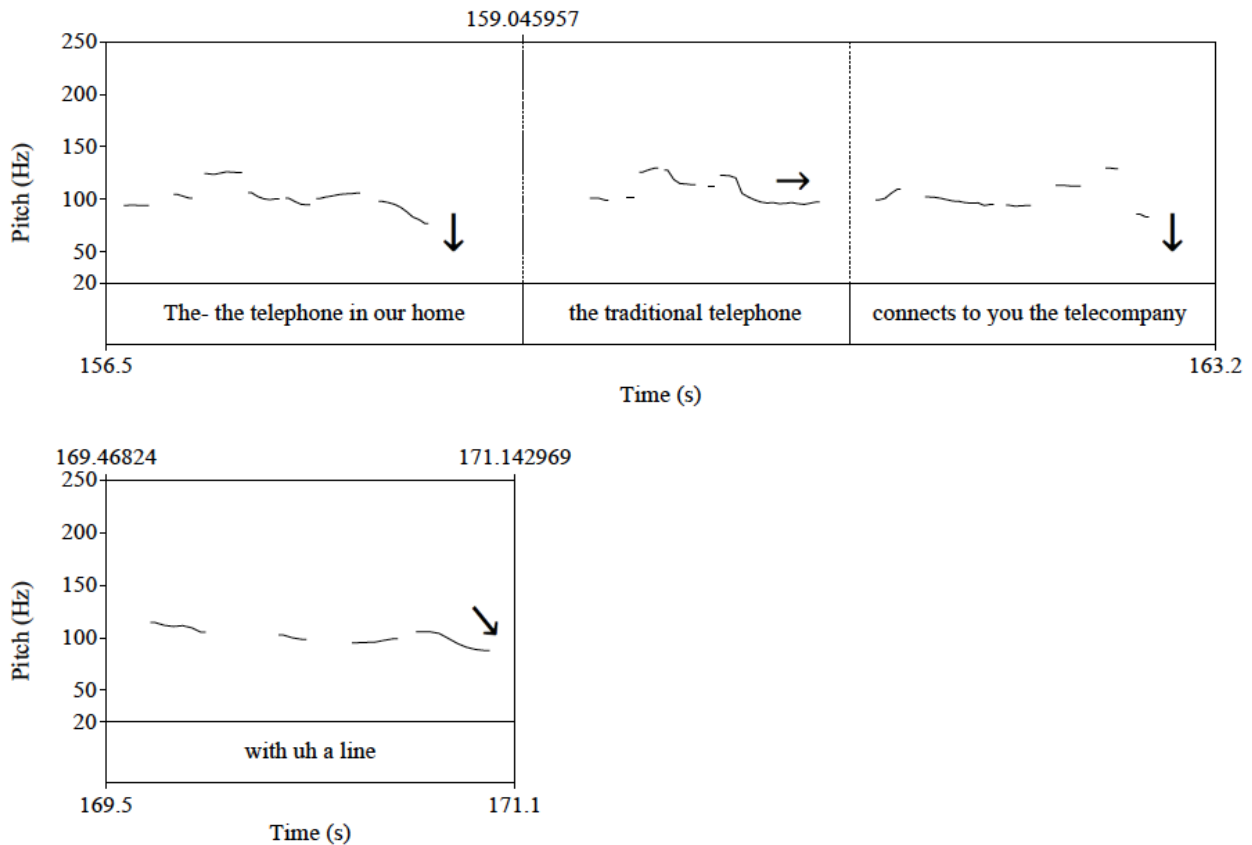

Figure 4. Pitch boundaries in Chen's first presentation.

DISCUSSION

Regarding the first research question, our findings are in line with the evidence that even advanced L2 speakers of English, and particularly L1 speakers of Mandarin, may fall short of native patterns of discourse intonation (Pickering, 2004; Wennerstrom, 1998). All of the L2 participants produced peaks at topic shifts that were considerably lower than those produced by the native controls. Less pronounced peaks at the beginning of new discourse topics might make the macro-structure of a presentation less salient to the audience and could possibly affect comprehensibility, given that previous research has shown a correlation between accurate use of prosody and comprehensibility ratings (Anderson-Hsieh, Johnson, & Koehler, 1992; Wennerstrom, 1998).

In terms of the second research question, the use of pitch at utterance-medial boundaries did not seem to be a challenge for the L2 speakers in this study, as their percentages of non-low boundaries approximated those of the native controls. This suggests that, overall, the participants were able to link *I*s intonationally, correctly signaling the dependency of mid-utterance *I*s on subsequent discourse for their interpretation. This finding differs from Wennerstrom (1998), where 10 out of 16 participants produced less than 80% of non-falling boundaries, possibly because they had lower oral proficiency in English.

The third and fourth research questions asked about potential changes across the two semesters of immersion. Longitudinal changes were observed for three of the four participants, particularly in their use of paratones. However, they did not change in the same way. The two female participants, who had the highest proficiency scores, increased their marking of topic shifts and their use of non-low boundaries over time. Yang showed a larger change in her paratones, as her percentages of pitch increase went up from P1 to P2 and from P3 to P4, whereas Xu plateaued after P2. Changes in utterance-medial pitch boundaries for both Xu and Yang were small due to the ceiling effect. Chen, one of the male participants, became less accurate in his use of paratones from P1 to P3, then improved considerably from P3 to P4, such that his end point was close to his starting point. Interestingly, the opposite happened to his pitch boundaries, which became more accurate over time until ceiling level was reached at P3. No changes were observed in Wei's use of paratones, and his utterance-medial boundaries were almost all non-low from P1, so there was no room for improvement.

Despite not having received any instruction on English pronunciation or academic speaking during the study, Xu and Yang might have improved naturalistically simply by being immersed in the L2 environment or by attending or giving academic presentations. The fact that they did better than the male participants might be due to possibly higher metalinguistic awareness, as previous research has found a positive relationship between language proficiency and awareness (Renou, 2001). Increased awareness may in turn facilitate further language development when the learners arrive in an environment with more opportunities for L2 input and output – and, consequently, more opportunities to notice the gap between the two. In contrast, a learner like Wei, who had lower English proficiency, might require more time in the L2 context, more practice or even explicit instruction in order to make progress.

The U-shaped curve observed for Chen's paratones has been previously reported in the literature for other L2 features, including verbs (Kellerman, 1979) and codas (Abrahamsson, 2003). U-shaped behaviour occurs when structures that are initially correct or closer to target "undergo a process of attrition, only to be reestablished at a later stage" (Sharwood Smith & Kellerman, 1989, p. 220). The learners' start and end point may be similar, but the process of cognitive restructuring that takes place between the two points changes their underlying knowledge of the language. This process might have been triggered when Chen moved to Canada and started to have more exposure to and practice with the language.

Given the pilot nature of this study, findings should be regarded as preliminary only, and no firm conclusions can be drawn at this point. It is still to be verified whether the other participants in the project will show similar development. Admittedly, the speech samples analyzed in this study do not necessarily reflect the speakers' spontaneous use of intonation, as the presentations could be prepared beforehand and were often aided by the use of PowerPoint and written notes. Also, the L2 presentations were not completely authentic in that the audience was student research assistants. This method was chosen, however, to make the presentations more

comparable and to allow participants' presentations to be recorded around the same time. Finally, another limitation was the analysis of only one presentation by each native speaker. Gathering similar longitudinal data from L1 speakers would allow for a better comparison with the L2 presentations and would help to control for task effects.

ABOUT THE AUTHORS

Larissa Buss (larissacks@gmail.com) is a PhD student in Education specializing in Applied Linguistics at Concordia University, in Montreal. Her research interests include teacher education, pronunciation instruction, and second language phonetics and phonology. She has a full scholarship for her doctoral studies from the CAPES Foundation of the Ministry of Education of Brazil.

Walcir Cardoso (Ph.D. McGill University, walcir@education.concordia.ca) is an associate professor of Applied Linguistics (Department of Education, Concordia University – Montreal, Canada). He conducts research on the second language acquisition of phonology, morphosyntax and vocabulary, and the effects of computer technology (particularly clickers, text-to-speech synthesizers, and automatic speech recognition) on L2 learning.

Sara Kennedy (sara.kennedy@education.concordia.ca) teaches at Concordia University in Montreal, Canada. She conducts research on the effects of instruction and exposure on listening and speaking in a second language, and on the production, comprehension, and assessment of L2 speech. Her current research projects involve immigrants and post-secondary students who are learning and studying in a second language.

REFERENCES

Abrahamsson, N. (2003). Development and recoverability of L2 codas: A longitudinal study of Chinese-Swedish interphonology. *Studies in Second Language Acquisition, 25*(3), 313–349.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*(4), 529–555.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer [Computer program]. Version 5.3.77, retrieved 18 May 2014 from http://www.praat.org/.

Brown, G. (1977). *Listening to spoken English*. London: Longman.

Chafe, W. (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Chaudron, C., & Richards, J. C. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics, 7*(2), 113–127.

Chun, D. M. (2002). *Discourse intonation in L2: From theory and research to practice*. Amsterdam and Philadelphia: John Benjamins.

Cruttenden, A. (1986). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.

Derwing, T. M., Thomson, R. I., & Munro, M. J. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System, 34*(2), 183–193.

Emonds, J. (1976). *A transformational approach to English syntax: Root, structure-preserving, and local transformations.* London: Academic Press.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84*(1), 70–9.

Grosz, B., & Hirschberg, J. (1992). Some intonational characteristics of discourse structure. *Proceedings of ICSLP* 92, Banff, 429–432.

Kellerman, E. (1979). The problem with difficulty. *Interlanguage Studies Bulletin, 4*, 27–48.

Lehiste, I. (1975). The phonetic structure of paragraphs. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception* (pp. 195–203). Berlin: Springer-Verlag.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology.* Dordrecht: Foris Publications.

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly, 35*(2), 233–255.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes, 23*(1), 19–43.

Renou, J. (2001). An examination of the relationship between metalinguistic awareness and second-language proficiency of adult learners of French. *Language Awareness, 10*(4), 248–267.

Sharwood Smith, M., & Kellerman, E. (1989). The interpretation of second language output. In H. W. Dechert & M. Raupach (Eds.), *Transfer in language production* (pp. 217–36). Norwood, NJ: Ablex.

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech, 37*(1), 21–43.

Thompson, S. E. (2003). Text-structuring metadiscourse, intonation and the signalling of organisation in academic lectures. *Journal of English for Academic Purposes, 2*(1), 5–20.

Tyler, A., & Davies, C. (1990). Cross-linguistic communication missteps. *Text & Talk, 10*, 385–411.

Tyler, A., Jefferies, A., & Davies, C. (1988). The effect of discourse structuring devices on listener perceptions of coherence in non-native university teacher's spoken discourse. *World Englishes, 7*(2), 101–110.

Wennerstrom, A. (1994). Intonational meaning in English discourse: A study of non-native speakers. *Applied Linguistics, 15*(4), 399–420.

Wennerstrom, A. (1998). Intonation as cohesion in academic discourse. *Studies in Second Language Acquisition, 20*, 1–25.

Wennerstrom, A. (2001). *The music of everyday speech.* New York: Oxford University Press.

Yule, G. (1980). Speakers' topics and major paratones. *Lingua, 52*, 33–47.