LONGITUDINAL CHANGES IN THE USE OF PARATONES IN L2 ENGLISH SPEECH BY MANDARIN SPEAKERS

MUDANÇAS LONGITUDINAIS NO USO DE PARÁGRAFOS FONOLÓGICOS EM INGLÊS (L2) POR FALANTES DE MANDARIM

Larissa Buss[1], Walcir Cardoso[2], Sara Kennedy[3]

Abstract: *First language (L1) English speakers have been observed to organize their oral discourse into macro-units analogous to paragraphs in writing. These units, called* paratones *(BROWN, 1977) or phonological paragraphs (TENCH, 1996; THOMPSON, 2003), are characterized by extra high pitch at the beginning of a new discourse topic (YULE, 1980). The present study investigated how seven second language (L2) graduate students' use of paratones developed naturalistically during their first six months immersed in an L2 environment. The participants, all L1 speakers of Mandarin, were recorded delivering four short academic presentations at approximately two-month intervals. Presentations given by two native English speakers were also analyzed for comparison. Overall, the L2 participants' pitch peaks at topic shifts were considerably less prominent than those observed in the native-speaker data. Only one participant's use of paratones seemed to change over time, showing improvement from the beginning to the end of the study.*
Keywords: pronunciation; intonation; paratone; L2 English.

Resumo: *Falantes de inglês como língua materna (L1) organizam seu discurso oral em macro-unidades semelhantes a parágrafos na escrita. Essas unidades, chamadas em inglês de* paratones *(BROWN, 1977) ou parágrafos fonológicos (TENCH, 1996; THOMPSON, 2003), são caracterizadas por um tom alto no início de um novo tópico discursivo (YULE, 1980). Este estudo investigou como o uso de parágrafos fonológicos desenvolveu-se naturalisticamente na fala de sete alunos de pós-graduação, todos falantes de inglês como segunda língua (L2), durante seus primeiros seis meses de imersão em um ambiente de língua inglesa. Os participantes, todos falantes de mandarim como L1, foram gravados realizando quatro pequenas apre-*

---

1 PhD Student, Applied Linguistics – Department of Education, Concordia University, Canada.
2 Professor at the Departament of Education - Concordia University, Canada.
3 Ph.D., Second Language Education, McGill University. Associate Professor – Concordia University., Canada.

sentações acadêmicas aproximadamente a cada dois meses. Apresentações realizadas por dois falantes de inglês como língua materna também foram analisadas para fins de comparação. Em geral, os picos de entonação produzidos pelos falantes de L2 nas mudanças de tópico foram consideravelmente menos salientes do que aqueles produzidos pelos falantes de L1. Apenas uma participante exibiu mudanças claras em seu uso de parágrafos fonológicos, demonstrando progresso ao longo do estudo.
Palavras-chave: pronúncia; entonação; parágrafos fonológicos; inglês como L2.

Intonation is an important cue to the organization of spoken discourse in L1 English. Native speakers use variations in pitch to demarcate different discourse topics, creating a macro-structure that is likely to aid listener comprehension. Tench (1996) observes that, in newsreading, it is intonation that helps listeners identify when one item of news has finished and another one begins. 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). 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). Another way of describing this phenomenon is to say that pitch range is expanded at the beginning of new topics, then lowered and compressed at the end (BRAZIL; COULTHARD; JOHNS, 1980; PIERREHUMBERT; HIRSCHBERG, 1990). 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* (TENCH, 1996; THOMPSON, 2003), *pitch sequences* (BRAZIL, 1997), *sequence chains* (BARR, 1990), and *intonational paragraphs* (LEVIS; PICKERING, 2004; PICKERING, 2004).

Although paratones have been extensively observed in L1 English, not much is known about how they are acquired and used by L2 speakers. This study analyzed the use of paratones in academic presentations given by seven L1 Mandarin/L2 English graduate students. Data were collected four times during their first six months of immersion in an L2 environment. The main purpose of the study was to examine potential changes in the target feature over time.

Paratones in L1 speech

A number of researchers have reported somewhat similar observations regarding the use of paratones to organize speech and the prosodic cues

associated with them. Swerts and Geluykens (1994) analyzed the use of paratones in L1 Dutch monologues. Their data consisted of spontaneous oral instructions on how to build a cardboard house from ready-made pieces, and each instruction was considered a new topic. The researchers compared the highest fundamental frequency (F_0) peak of each topic-introducing clause (called "topic peaks") to the highest F_0 peaks of the following clauses (called "non-topic peaks"). The findings revealed that the speakers tended to make topic peaks more prominent than non-topic peaks by producing them with a higher pitch.

Thompson (2003) analyzed the use of phonological paragraphs in authentic and pedagogically prepared academic lectures in English. Phonological paragraphs were identified impressionistically by two independent raters, who listened for "low pitch on the final tonic syllable of the paragraph followed by a jump up to high pitch on the first prominent syllable of the new paragraph" (p. 9). The raters also attended to whether this pitch was exceptionally high when compared with other paragraphs and whether the previous paragraph ended with extra low pitch, decreased volume or speed, laryngealization (i.e., creaky voice), and/or a long pause. The study found that phonological paragraphs were used in both types of lectures, although they were longer in authentic discourse.

Using discourse segments as their units of analysis, Grosz and Hirschberg (1992) found evidence that listeners pick up on these discourse-structuring intonational cues. They defined discourse segments as groups of utterances that contribute to the same underlying purpose or intention of the speaker/writer (e.g., inform about how people reacted to an accident). In one of their pilot studies, seven listeners labeled the discourse structure of a news story, either from text alone (with most of the 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 that used read-aloud speech, Lehiste (1975) had a native English speaker record six three-sentence paragraphs that consisted of different orderings of the same three sentences, all of which seemed equally possible. The sentences were then isolated, randomized, and presented to 30 native listeners, who had to determine their context of production (isolated, paragraph-initial, medial or final). The results revealed that sentences produced with high fundamental frequency 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 phonological paragraph.

Paratones in L2 speech

In L2 speech, listener expectations regarding the use of discourse intonation are not always met. Wennerstrom (1994) analyzed paratones and several other features in the speech of 10 L1 and 30 L2 English speakers from three language groups: Spanish, Japanese, and Thai. The findings revealed consistency among native speakers in their use of intonation, but several differences across L1s. Overall, the Spanish group shared the most similarities with the native English speakers. Japanese and Thai subjects, on the other hand, did not increase pitch range to signal new topics like L1 English and Spanish speakers did.

Pickering (2004) compared how American and Chinese teaching assistants (TAs) used intonational paragraphs in their teaching discourse in English. Results showed that native-speaking TAs consistently used pitch to structure their speech into coherent sections. Conversely, although the Chinese TAs produced the phonological cues associated with intonational paragraphs, these cues did not necessarily correspond to semantic or structural boundaries in their discourse.

There is some evidence that inaccurate use of paratones may undermine effective communication. Wennerstrom (1998) analyzed the speech of 18 Mandarin speakers attending graduate programs in the United States. All of them had taken a 10-week English course focused on developing their pronunciation, presentation skills, and teaching techniques. The data for the study came from the participants' final exam for this course and consisted of short lectures in their fields of study. Three raters scored the exams on different categories, including "production", which was primarily based on level of comprehensibility. In analyzing the use of paratones in the lectures, Wennerstrom found a significant positive relationship between paratone accuracy and the production scores the speakers received on the exam.

Despite the prevalence of paratones in L1 English, little is known about how they develop in the speech of L2 learners immersed in an English-speaking environment. Previous research has suggested that some L2 learners can improve their pronunciation even without explicit instruction in this area. However, most of the improvement seems to take place within the learners' first year of immersion in the L2 environment, provided that they have frequent exposure and interactions in the L2 (DERWING; MUNRO; THOMSON, 2007; DERWING; THOMSON; MUNRO, 2006; DERWING; MUNRO, 2013; FLEGE, 1988).

The current study offers a contribution to our understanding of whether and how the use of paratones develops naturalistically in L2 English speakers during their first six months immersed in an L2 context. The genre analyzed was the academic presentation, and it was chosen for two main reasons. First, oral presentations are an important part of graduate education, as graduate students are usually expected to give presentations in class and at conferences. In addition, they are a form of monologic discourse, which facilitates the analysis of a macro-unit like the paratone.

Research questions

The research questions addressed by this study are:

1. Do L2 English/L1 Mandarin speakers differ from L1 English speakers in their use of paratones? If so, how?
2. Does the L2 speakers' use of paratones change across two terms of immersion in an academic, English-speaking environment? If so, how?

Methods

Participants and data collection

The participants were seven L1 speakers of Mandarin (three males and four females), all in their twenties, studying in graduate programs at a Canadian University. One of them was in computer science, another was in instructional technology, and the rest were in engineering programs. All of the participants had arrived in Canada approximately one month before the beginning of the study and had never lived abroad before. They had a minimum of 6.5 on each band of the International English Language Testing System (IELTS) or a minimum score of 80 on the Internet-based Test of English as a Foreign Language (TOEFL), which was the entry level required by the university at the time. They did not receive any formal instruction on English pronunciation or academic speaking during the course of the study. Two native English speakers from Canada (one male and one female) also participated as controls. They were graduate students at the same university and were unaware of the purpose of the study.

Speech samples from the L2 graduate students were collected four times over their first two terms of study, at approximately two-month intervals (total length = 26 weeks). The two native controls were recorded only twice: at the beginning and at the end of their first term. 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 fields to an imagined audience of first-year undergraduate students.

Coding of discourse topics

The presentations were transcribed and coded for discourse topics. 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) or, quite simply, as "what is generally being talked about" (YULE, 1980, p. 33). Coding was also aided by the presence of "macro-markers" (CHAUDRON; RICHARDS, 1986), such as "The next thing is about…" and "The third principle is that…". When used, these types of 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 (CHAUDRON; RICHARDS, 1986), such as "Well" and "Also".

Phonetic cues, such as long pauses and higher pitch peaks, usually coincided with new discourse topics and corroborated the coding. Although we did not rely on these cues to identify the discourse topics in the presentations, they were taken into consideration in order to establish the exact place where one topic ended and another began. In some presentations, there was more than one place where the boundary between two topics could be placed.

For example, by reading the transcript of Participant 7's second presentation (P2), it was easily observed that the second discourse topic was on the features and utility of smart materials, while the third topic was on the hysteresis effect, a problem found in smart materials. However, without listening to the recording, it was not completely clear whether topic 3 would begin after (2) or in (3) in the excerpt below.

> (1) Therefore, they can be utilized in the areas of nanotechnology and mechanical engineering and uh aerospace. (2) So uh but uh most of smart materials has the hysteresis effect. (3) The term-the word hysteresis comes from the ancient Greek, which means lag behind in smart materials… (Participant 7, P2)

The audio revealed a long pause before (2) and a higher F_0 peak on the word *most*, whereas (3) was produced with no particular emphasis. Thus, (2) was taken to be the beginning of topic 3. It is important to note that even when L2 speakers did not produce a particularly high pitch increase (as did the L1 controls), there was almost always some level of increase at topic shifts. Nonetheless, as already mentioned, prosodic criteria were only used to determine the exact start point of a topic that had already been identified through the other criteria explained above (i.e., semantic content and discourse markers), whenever the transcript revealed more than one possible start point. Prosodic cues were not used to identify the topics themselves, in order to avoid circularity.

Brown, Currie, and Kenworthy (1980) warned against the danger of circularity and the need for an independent theory of topic structure. Certainly, as also noted by Swerts and Geluykens (1994), the role of prosody in demarcating topic structure can only be properly investigated if one is able to establish this structure *before* analyzing prosody. This is especially true when studying nonnative speech, as L2 speakers will not always produce the intonation patterns found in L1 speech. Thus, in order to capture their difficulties with these intonation patterns (and eventual changes over time or after instruction), it is important to have an *a priori* idea of the places in the discourse where these patterns are expected to occur.

Data analysis

As previously mentioned, the paratone is characterized by a high peak in the first prominent syllable of a new discourse topic, followed by less prominent peaks (SWERTS; GELUYKENS, 1994; YULE, 1980). Thus, there is a jump up to a higher pitch that takes place at topic shifts (THOMPSON, 2003). Based on this description, the analysis of the paratone was done by measuring on Praat (BOERSMA; WEENINK, 2014) the first F_0 peak in each topic-initial and topic-final intonational phrase (see below for a description of the intonational phrase). Then, the F_0 value found for the final phrase of one discourse topic was subtracted from the F_0 value found for the initial phrase of the subsequent topic. This resulted in measures of pitch increase at topic shifts. Each participant's measures were converted into percentages and averaged.

The *intonational phrase* (NESPOR; VOGEL, 1986), also called *intonation unit* (CHAFE, 1994; HIMMELMANN, 2006), *tone unit*

(BRAZIL, 1997), *intonation group* (CRUTTENDEN, 1986), and *intermediate phrase* (PIERREHUMBERT; HIRSCHBERG, 1990)*,* is the domain of an intonation contour. Its boundaries were identified by a series of phonetic cues listed in the literature. These cues include accelerated delivery at the beginning of a new intonational phrase (i.e., anacrusis); lengthening, low intensity and sometimes laryngealization (i.e., creaky voice) at the end of a phrase; a change in the pitch of the final unaccented syllable; a jump in pitch from the end of one phrase to the beginning of the next; and a pause between two phrases (CHAFE, 1994; CRUTTENDEN, 1986; HIMMELMANN, 2006).

Naturally, not all of these cues need to be found at every boundary, and sometimes only one of them will be present. However, it is important to note that pauses alone do not always denote a boundary. As explained by Himmelmann (2006), pauses occurring at intonational phrase boundaries are usually unfilled and characterized by an audible relaxation of the speaker's vocal organs, audible exhalation, and/or audible inhalation. On the other hand, *hesitation pauses* tend to be rather abrupt interruptions of speech, often ending in a glottal stop, and include fillers (*uhm, uh*) and further disfluencies after them (HIMMELMANN, 2006). These aspects were taken into consideration in identifying the boundaries of topic-initial and topic-final intonational phrases. The boundary cues were identified based on auditory impression and observations of the F_0 contour on Praat.

Results

The number of discourse topics (or paratones) identified in each presentation varied between two and eight (mean = 5.1) and totaled 169 in all presentations. The analysis of F_0 peaks at the beginning of topic-initial and topic-final intonational phrases revealed higher percentages of pitch increase for the L1 controls than for L2 speakers. As shown in Figures 1 and 2, the native speakers increased their pitch, on average, by 50.4% (Female) and 67.2% (Male) to signal topic shifts, whereas the nonnative speakers' mean increases ranged from 6.5% (Participant 1, P1) to 42.8% (Participant 6, P4), with a total average of only 23.6%. Only Participant 6 had mean increase values around 40%, while those of the other participants did not exceed 31.8%. Overall, both the male and the female participants were well below their L1 counterparts in their marking of topic shifts.
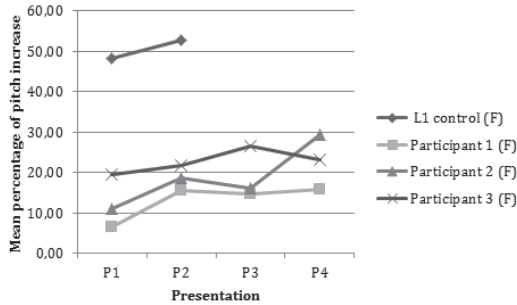
Figure 1. Mean percentage of pitch increase at topic shifts for female participants.
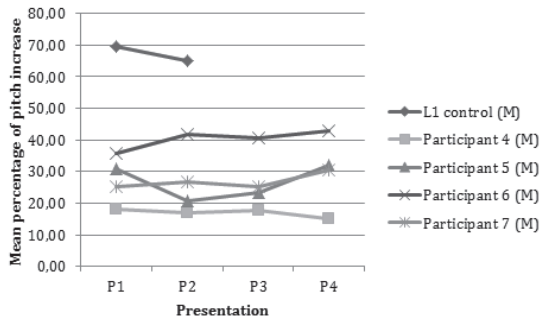

Figure 2. Mean percentage of pitch increase at topic shifts for male participants.

Figure 3 shows an example of a large pitch increase from the last intonational phrase of topic 2 to the first intonational phrase of topic 3 in Participant 6's fourth presentation. The prominent syllables corresponding to the F_0 peaks are indicated in capital letters, and the pitch values measured are shown in brackets.
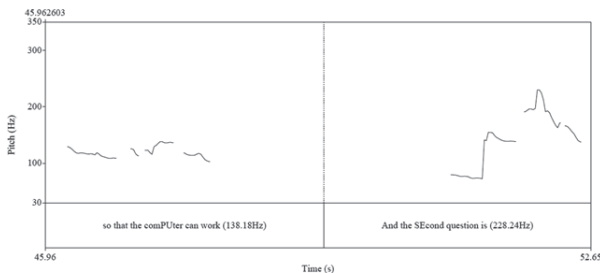

Figure 3. Pitch increase at a topic shift in Participant 6's fourth presentation.

There was considerable variability in pitch increase values at topic shifts, even within a single presentation. One of the most striking examples of this is Participant 4's second presentation, where the transition from topic 1 to topic 2 was marked by a 20.5% pitch increase, but the beginning of topic 3 had a mere 4.2% increase. From topic 3 to 4, there was a 10.8% increase, then a 31% increase from topic 4 to 5. Variability was also found in the native speaker data, possibly indicating different levels of semantic disjunction. Nonetheless, native speakers seemed to have a threshold level of pitch increase that was necessary for adequate marking of new topics. They would never start a new topic without a minimum 30% pitch increase, and they sometimes increased their pitch in 80% and more, something that the L2 speakers never did.

Longitudinally, the L1 speakers' mean percentage of pitch increase did not change much from their first to their second presentation. The female control produced an average of 48.1% increase at P1 and 52.7% at P2, whereas the male control produced 69.6% and 64.9%. Thus, the difference between the two presentations given by the L1 controls was approximately 5%. It was in fact expected that no considerable differences would be observed in the native speaker data, given that the target feature is not likely to be developing in their speech.

As for the L2 data, most of the participants did not seem to exhibit considerable changes in their marking of topic shifts, as shown in Figures 1 and 2. Although there were differences from one presentation to another, most of them did not exceed by much the 5% difference found in the native speaker data. Therefore, we would be reluctant to suggest that they reflect actual linguistic development. Yet one of the female participants (Participant 2) did show considerable change. She started the study with a very low percentage of pitch increase at P1 (only 11%), but seemed to improve over time, reaching an average of 29.4% in her last presentation (see Figure 1).

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 pitch 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).

Regarding the second research question, only Participant 2 seemed to have improved in her use of paratones over time. Despite not having received any instruction on English pronunciation or academic speaking during the study, Participant 2 might have improved naturalistically, simply by being immersed in the L2 environment or by attending or giving academic presentations. There is, in fact, previous evidence that L2 pronunciation can improve even in the absence of instruction, especially during the first year of immersion (DERWING *et al.*, 2006; FLEGE, 1988).

On the other hand, clear longitudinal development was not observed for the other participants, whose small changes are believed to reflect normal variability (as that found for the L1 controls). There are many potential explanations for this result. First, contrary to the participants in Derwing *et al.* (2006), the L2 speakers in this study were not beginners when they moved to Canada. They had all studied English in China for many years and were accepted to graduate programs in an English medium university. Thus, it is possible that this aspect of their pronunciation fossilized and would require explicit instruction in order to change. Another possibility is that not all features of pronunciation are equally prone to naturalistic acquisition. Paratones may be more difficult for L2 learners to acquire spontaneously, perhaps because people are generally unaware of this use of intonation and "noticing" (SCHMIDT, 1990) does not take place. Alternatively, the participants might not have had the necessary amount of exposure or practice in the L2 to trigger spontaneous improvement. A further hypothesis is that six months may not have been enough time to observe development. If the study had run for an entire year, maybe changes in more participants would have been observed.

The results reported in this study should be interpreted considering the following limitations. To begin with, the speech samples analyzed 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. In fact, some of the participants relied rather heavily on their slides and notes. Also, the 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, some presentations had a small number of discourse topics (only two, three or four), which limited the number of F_0 measures taken for these presentations. Given that considerable variability was found within the presentations, a larger number of measures would have been desirable. In future studies, this issue could be addressed by having participants give longer presentations or more than one short presentation at each data collection point.

REFERENCES

ANDERSON-HSIEH, J.; JOHNSON, R.; KOEHLER, K. The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*(4), 529–555, 1992.
BOERSMA, P.; WEENINK, D. Praat: software livre. Versão 5.3.77. Amsterdam: 2014 Disponível em: <http://www.praat.org>. Acesso em: 18 maio 2014.
BRAZIL, D. *The communicative value of intonation in English*. New York: Cambridge University Press, 1997.
BRAZIL, D.; COULTHARD, M.; JOHNS, C. *Discourse intonation and language teaching*. London: Longman, 1980.
BROWN, G. *Listening to spoken English*. London: Longman, 1977.
BROWN, G.; CURRIE, K.; KENWORTHY, J. *Questions of intonation*. London: Croom Helm, 1980.
CHAFE, W. *Discourse, consciousness, and time*. Chicago: University of Chicago Press, 1994.
CHAUDRON, C.; RICHARDS, J. C. The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, *7*(2), 113–127, 1986.
CRUTTENDEN, A. *Intonation* (2nd ed.). Cambridge: Cambridge University Press, 1986.
DERWING, T. M.; MUNRO, M. J. The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, *63*(2), 163–185, 2013.
DERWING, T. M.; MUNRO, M. J.; THOMSON, R. I. A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, *29*(3), 359–380, 2007.
DERWING, T. M.; THOMSON, R. I.; MUNRO, M. J. English pronunciation and fluency development in Mandarin and Slavic speakers. *System*, *34*(2), 183–193, 2006.

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

GROSZ, B.; HIRSCHBERG, J. Some intonational characteristics of discourse structure. *Proceedings of ICSLP 92*, Banff, 429–432, 1992.

HIMMELMANN, N. P. The challenges of segmenting spoken language. In J. Gippert, N. P. Himmelmann, & U. Mosel (Eds.), *Essentials of language documentation* (pp. 253–274). Berlin: Mouton de Gruyter, 2006.

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

LEVIS, J., & PICKERING, L. Teaching intonation in discourse using speech visualization technology. *System*, *32*(4), 505–524, 2004.

NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris Publications, 1986.

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

PIERREHUMBERT, J.; HIRSCHBERG, J. The meaning of intonational contours in discourse. In *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press, 1990.

SCHMIDT, R. The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158, 1990.

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

TENCH, P. *The intonation systems of English*. London: Cassell, 1996.

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

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

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

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