BUCLD 43 Proceedings
To be published in 2019 by Cascadilla Press
Rights forms signed by all authors

L2 Adaptation to Unreliable Prosody During Structural Analysis: A Visual World Study

Chie Nakamura[1], Jesse A. Harris[2], Sun-Ah Jun[3], and Yuki Hirose[4]

1. Introduction

In order to comprehend spoken language, listeners need to analyze the speech signal according to the language-specific prosodic structure which is typically realized with pitch (intonation), intensity, and duration. The edge of a major prosodic unit is known to align with the edge of a major syntactic unit, and is often marked by a boundary tone, phrase-final lengthening, and a pause (Beckman & Pierrehumbert 1986, Selkirk, 2000, 2007, 2011; Nespor and Vogel 1986; Truckenbrodt 1999, Shattuck-Hufnagel & Turk 1996, Jun 2005). Past research on first language (L1) processing has shown that listeners use prosody to resolve syntactic ambiguity in structural analysis. Accumulating evidence in L1 processing studies suggests that prosodic information is considered during very early processing stages. For instance, Snedeker and Trueswell (2003) examined globally ambiguous sentences such as (1), which is ambiguous as to whether the prepositional phrase (PP) *with the flower* is attached to the VP (i.e., tap the frog by using the flower, Instrument interpretation) or it is attached to the NP (i.e., tap the frog that has the flower, Modifier interpretation). In a series of experiments, they investigated how speakers produce, and how listeners use, prosodic cues to distinguish alternative meanings of the sentence.

(1) Tap the frog with the flower.

Their production results showed that speakers produced different prosodic cues only when they were aware of the ambiguity. When speakers were aware of the alternative meaning of the sentence, they tended to lengthen the direct object noun (*frog*) and located a prosodic boundary between the noun and the ambiguous PP (*with the flower*) for the instrument interpretation. In contrast, they tended to lengthen the verb (*tap*) and locate a prosodic boundary after the verb for the modifier interpretation. They also observed that prosodic cues produced by speakers influenced listeners' preferential looking at targets in a

[1] Chie Nakamura, University of California, Los Angeles /JSPS, chienak@mit.edu
[2] Jesse A. Harris, University of California, Los Angeles
[3] Sun-Ah Jun, University of California, Los Angeles
[4] Yuki Hirose, University of Tokyo

visual world paradigm study. Importantly, the effects were observed shortly after the onset of the direct object noun (*the frog*) and prior to the onset of the ambiguously attached preposition (*with*). Their results demonstrate that listeners rapidly integrate prosodic information to determine the syntactic structure of lexical input, and that prosodic cues lead listeners to make predictions about information that has yet to be heard.

In line with Snedeker & Trueswell (2003), there are a growing number of studies reporting that listeners process prosodic cues rapidly enough to *anticipate* a likely speaker-intended referent. It has been shown that listeners make expectations before they encounter disambiguating lexical material in interpretations of discourse and information structure (Dahan, Tanenhaus, & Chambers, 2002; Ito & Speer, 2008) and in structural prediction in different languages (Nakamura, Arai, Mazuka, 2012 in Japanese; Weber, Braus, & Crokcker, 2006 in German). These studies provide evidence that prosodic information is integrated incrementally in online processing, supporting the view that language users generate real-time expectations about upcoming linguistic input using prosodic cues in conjunction with other information such as discourse context and world knowledge.

Still, relatively little is known about how prosodic cues influence the interpretation and comprehension in L2 sentence processing. In L1 processing research, prosodic boundary information is known to have a strong effect on syntactic ambiguity resolution in processing sentences like (2) (Schafer, Speer, Warren, & White, 2000). In (2), "the square" is temporarily ambiguous as to whether it serves as the subject of the following verb, *will* (i.e., *early closure*), or the direct object of the preceding verb, *moves* (i.e., *late closure*). Hwang & Schafer (2006) showed that L2 learners chose an early closure continuation for temporarily ambiguous segments more often when the first prosodic boundary (marked as %) was stronger relative to the second boundary, and vice versa.

(2) a. *Early closure:*
When that moves % the square % will encounter a cookie.
(2) b. *Late closure:*
When that moves % the square % it'll encounter a cookie.

Their study suggests that, like native speakers, L2 learners are sensitive to prosody-syntax alignment in building structural representations (see also Nickels & Steinhauer 2016 for similar results with Chinese and German learners of English L2 using event-related potentials).

However, other studies that investigated L2 sentence processing using online measures have argued either L2 learners are not able to make use of linguistic information as fast as native speakers (Ito, Pickering, & Corley, 2018; Martin, Thierry, Kuipers, Boutonnet, Foucart, & Costa, 2013), or that L2 learners have processing limitations in anticipating upcoming information (Grüter & Rohde, 2013). These studies suggest that L2 learners' ability to use linguistic information to predict incoming input is weaker compared to native speakers. L2 learners' weaker ability might therefore delay or even prevent them from making detailed predictions about upcoming linguistic material. The results raise the further possibility that even though L2 learners are sensitive to prosodic boundary information in general, they are less sensitive to the prosody-syntax mapping during incremental processing, and, as a consequence, do not generate strong structural predictions on the basis of prosodic boundary information.

Recent work on L1 sentence processing has also addressed how processing expectations change with varying linguistic input, a process sometimes known as *linguistic adaptation* (e.g., Norris, McQueen, & Cutler, 2015; Fine, Jaeger, Farmer, & Qian, 2013). These studies propose that comprehenders use statistical regularities to make predictions about upcoming information, and that they update distributional statistics when the input differs from what was expected. These accounts straightforwardly explain how language users accommodate linguistic variability; expectations are continuously modified during exposure. The view that expectations for future input are constantly updated according to what participants experience within the experiment points us to another interesting possibility in L2 processing. If adaptation is driven by statistical linguistic knowledge and L2 speakers have weaker access to distributional regularities, it is possible that they may be less able to rapidly adapt to variability in the linguistic input, compared to L1 processing.

The present study tested two questions. First, how prosodic boundary information influences the attachment of syntactic phrases in globally ambiguous sentences in L2 processing? Second, do L2 learners adjust the degree to which they use prosodic information when prosodic boundaries do not align with syntax? We also manipulated the plausibility of the sentence final word as an instrument to test whether L2 learners predict syntactic structure on the basis of prosodic information. If so, L2 learners should show processing difficulties arising from a mismatch between prosody and lexical information.

2. Experiments

The current study examines the impact of a prosodic boundary placed either before or after the patient NP (*the tiger*) in (3). Sentence (3) is structurally ambiguous as to whether the prepositional phrase (PP) modifies the noun (i.e., as an NP modifier: *The boy will see the tiger* that has *the binoculars*) or the verb (i.e., as an instrument of the verb: *The boy will see the tiger* by using *the binoculars*). In addition to the location of the prosodic boundary (%), the final word of the sentence was also manipulated to be either plausible (*binoculars*) or implausible (*popcorn*) with the instrument interpretation (4). This manipulation was included to test the prediction that a noun that is implausible as an instrument would elicit a processing penalty only when paired with an instrument prosody, which can be taken as evidence that listeners use prosodic boundary information to anticipate the syntactic structure.

(3) *Plausible instrument condition*

a. *Modifier prosody:*

The boy$_{L\text{-}H}$ % will see$_{L\text{-}H}$ % the tiger with the binoculars.

b. *Instrument prosody:*

The boy$_{L\text{-}H}$ % will see the tiger$_{L\text{-}H}$ % with the binoculars.

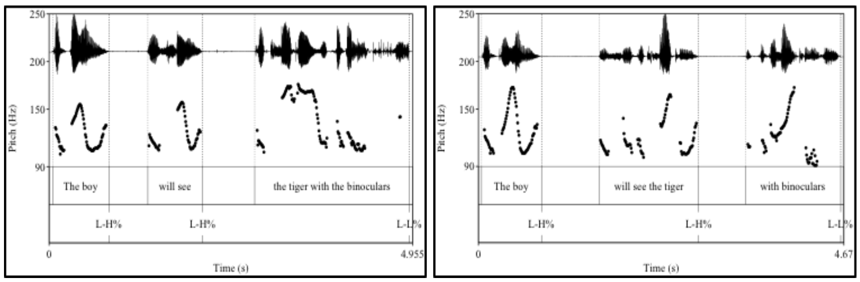


Fig. 1 Wave form and pitch track of the sentence in (3), labeled with boundary tones for (3a, left) and (3b, right)

(4) *Implausible instrument condition*

a. *Modifier prosody:*

The boy$_{L\text{-}H}$ % will see$_{L\text{-}H}$ % the tiger with the popcorn.

b. *Instrument prosody:*

The boy$_{L\text{-}H}$ % will see the tiger$_{L\text{-}H}$ % with the popcorn.

In addition, we manipulated the filler items presented with the experimental items in a between-subjects design. In Experiment 1, filler items had normal prosody such as (5a). In Experiment 2, filler items had prosody that did not provide reliable cues to the syntax, thus 'uninformative' prosody, such as (5b), where a falling L-L% boundary tone was produced between the determiner and the sentence-final NP. The L-L% pattern is typically used at the end of the sentence and a prosodic boundary is typically not present between a determiner and a noun, and would therefore be deviant or uninformative at the location shown in (5b).

(5) a. *Normal filler prosody (*Experiment 1)

The boy$_{L\text{-}H}$ % will touch$_{L\text{-}H}$ % the necktie and the razor.

b. *Uninformative prosody* (Experiment 2)

The boy$_{L\text{-}H}$ % will touch the necktie and the$_{L\text{-}L}$ % razor.

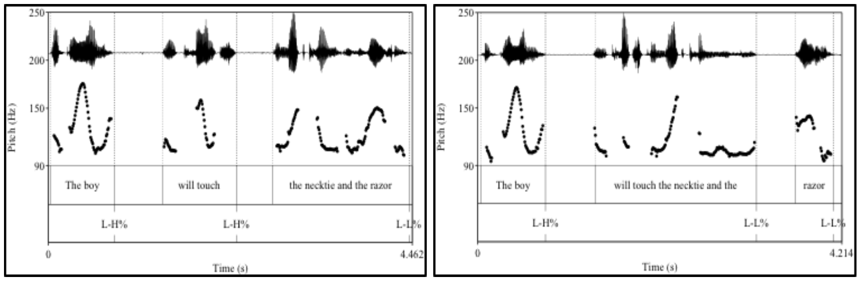
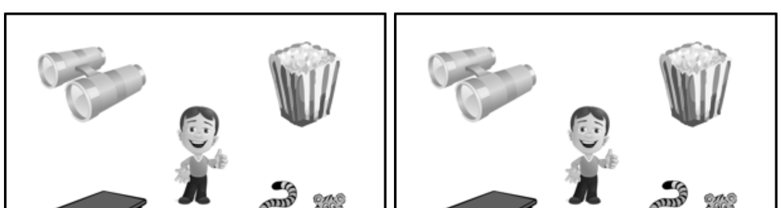


Fig. 2 Normal filler prosody (5a, left) and uninformative filler prosody (5b, right)

If prosody has an early effect in L2 sentence processing, listeners may adopt a structural analysis based on the location of prosodic boundaries before they encounter the critical word. In addition, if listeners adjust the degree to which they use prosodic cues according to how informative they are in online processing, it is predicted that listeners utilize prosodic information less when prosodic boundaries do not always align with the syntax.

Twenty-four experimental items were created in Experiment 1 and 2. Each item consisted of an auditory sentence and a corresponding visual scene. The auditory stimuli were recorded by a native speaker of English. The average speech rate (articulation rate, calculated by the number of syllables divided by total time minus pausing time) was 3.97 syllables/sec (De Jong & Wempe, 2009).

The visual scenes were prepared using clipart images. The position of objects was counter-balanced across the pictures. Four experiential lists were created following a Latin square design including 48 fillers. Each list had 72 items, presented in pseudo-random order. To keep participants focused on the task, a comprehension question appeared after each trial and participants were asked to click on the correct answer (e.g., "Which of the word was NOT mentioned in the sentence?" "binoculars", "spider", and "tiger" for (3)).



a. Visual scene presented with (3a,b) b. Visual scene presented with (4a,b)

Fig. 3 Visual scenes presented in (a) the plausible instrument condition and (b) the implausible instrument condition

Participants
Native Japanese speakers learning English were recruited from the student community at the University of Tokyo. They were paid in cash for their participation. All participants self-reported normal visual acuity and hearing. Thirty-two L1-Japanese L2-English speakers participated in Experiment 1, and 28 participated in Experiment 2. In addition, 32 native English speakers participated in each experiment as a control group. The standardized proficiency test score of our L2 participants corresponded to the proficiency levels of Intermediate to Advanced in the Common European Framework of Reference for Languages (CEFR).

Procedure
After participants received experimental instructions, we conducted a 9-point calibration procedure. They were told to listen to the auditory sentences carefully while paying attention to the picture on the computer monitor, and to correctly answer the comprehension question following each sentence. The auditory stimuli were presented via loud speakers. In each trial, an auditory sentence was presented 2500ms after the picture onset. Participants' eye-movements on the picture was recorded with Tobii TX300 at sampling rate of 300 Hz. The whole experimental session took approximately 30 minutes.

3. Data analysis and results

The fixation coordinate from the eye tracker was mapped onto five entities in the visual scene and then converted into gaze location to a visual target. We manually marked the onset of the ambiguous PP (*with the binoculars/popcorn*) and the noun within the PP (*binoculars/popcorn*) in each target sentence. For the analysis, we summed the proportion of looks to each object in the scene and calculated the logit of looks to each entity out of looks to all the objects in the scene, including background (Barr, 2008). We then conducted statistical analyses using linear mixed effects regression models (Baayen, Davidson, & Bates, 2008). Below, we first report a brief summary of the results of the native speakers, followed by a full report of the results of L2 learners.

3.1. Native speakers (Control group)

In Experiment 1, native English speakers looked significantly more at the modifier object (the tiger holding binoculars/popcorn in Fig. 3) when presented with Modifier prosody than with Instrument prosody. In addition, they looked more at the instrument object (the instrument binoculars in Fig. 3) when presented with Instrument prosody than with Modifier prosody. Crucially, the effects were observed before participants heard the final word, demonstrating that native speakers use prosodic boundary information in predictive structural processing. Also, the analysis of the proportion of looks to the instrument object for the duration of the final word (*binoculars/popcorn*) showed that the difference between the two types of prosody was significant only in the implausible instrument condition. This suggests that native speakers anticipated an instrument noun when presented with a prosodic boundary consistent with the Instrument interpretation, and experienced processing difficulty upon hearing an implausible instrument.

In Experiment 2, in which filler items had "uninformative" prosody, an anticipatory effect was observed only for the proportion of looks to the modifier object, and not to the instrument object. In addition, an analysis of how the anticipatory looks to the modifier object changed as the experiment progressed revealed that the difference between the two types of prosody was significant only in the first half of the experiment. The results together support the claim that native English speakers use prosodic boundary information in predicting upcoming information. The results further indicate that native speakers also evaluate and track how informative prosodic cues are and adjust the extent to which they use prosody to guide structure building processes (see Nakamura, Harris, & Jun (2019) for full report on the results of native speakers).

3.2. L2 learners

3.2.1. Experiment 1 (Normal filler prosody)

With L2 learners, no effects of interest were found in the anticipatory time window (i.e., before the onset of the sentence final word). Below, we report the analysis for the duration of the sentence final word (*binoculars/popcorn*). Prosody (Modifier or Instrument prosody), Noun Type (Plausible instrument or Implausible instrument), the interaction between Prosody and Noun Type, and Trial Order (1 to 72) were included as fixed factors in the model. Participants and items were included as random slopes. We computed the best-fitting model for each analysis using a backward selection approach by eliminating non-significant predictors from a fully specified model. For each analysis, we report coefficients, standard errors, t-value, and their *p*-values from the best-fitting model.

We first analyzed the proportion of looks made to the modifier object (the tiger holding binoculars/popcorn in Fig. 3). If the location of the prosodic boundary influences which structure is preferred, then participants should look more at the modifier object when the sentence has Modifier prosody than when it has Instrument prosody. The analysis was conducted from the onset of the sentence final word in each item until the mean offset of the sentence (1110ms). Fig. 4a shows the proportion of looks to the modifier object from the onset of the sentence final word until the mean sentence offset. Table 1 shows the results from the best-fitting model. There was a main effect of Prosody ($p<0.01$), as participants looked at the modifier object significantly more when presented with Modifier prosody than with Instrument prosody.

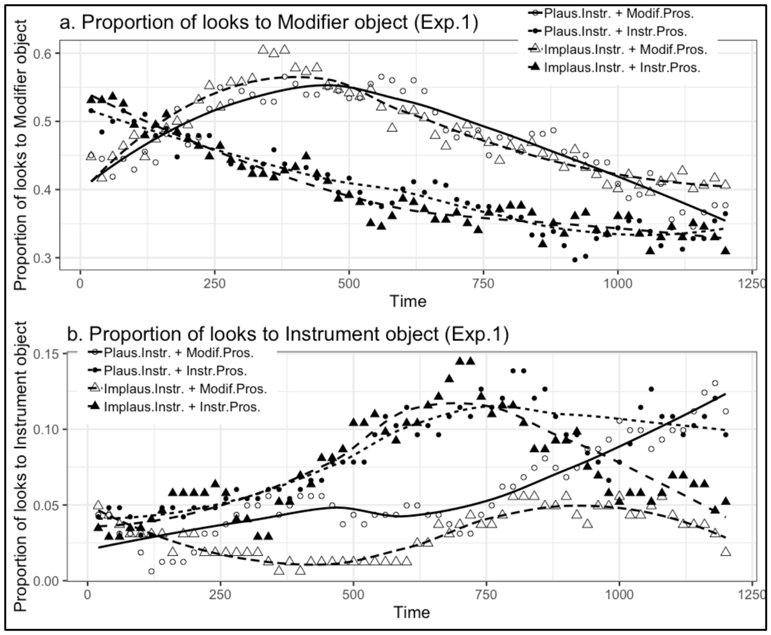


Fig. 4 Proportion of looks to (a) the modifier object and (b) instrument object in Experiment 1

Table 1. Analysis of looks to the modifier object for the duration of the final word

	Estimate	*Std. Error*	*t*-value	*p*-estimate
Intercept	1.68	0.33	5.10	<.001
Prosody	-9.45	0.28	-3.32	<.001
Noun Type	-0.11	0.28	-0.38	0.701
Trial	0.001	0.001	0.88	0.388
Prosody*Noun Type	-0.19	0.57	-0.33	0.741

Next, we analyzed the looks made to the object compatible with an instrument interpretation (the instrument binoculars in Fig. 3) for the same duration from the onset of the sentence final word. If L2 learners use prosodic boundary information to guide early structure building processes, we expect to observe more looks to the instrument object with Instrument prosody than with Modifier prosody. Additionally, they would experience processing difficulty in the Instrument prosody + Implausible instrument condition (4b).

Fig. 4b shows the proportion of looks to the instrument object from the onset of the final word until the mean sentence offset. Table 2 provides the results from the best-fitting model. There was a main effect of Prosody ($p<0.01$), so that participants looked at the instrument object significantly more with Instrument prosody than with Modifier prosody. There was also a main effect of

Noun Type (p<0.01), so that participants looked more at the instrument object in the Plausible instrument condition than in the Implausible instrument condition. This is most likely due to the visual scene presented in the Plausible instrument condition. Participants looked at the instrument object more often when there were two of the same instrument objects in the visual scene (e.g., two binoculars, Fig. 3a), regardless of prosody type.

The analysis between prosody types in the Implausible instrument condition revealed that participants looked significantly more at the instrument object in the Instrument prosody + Implausible instrument condition than in the Modifier prosody + Implausible instrument condition (p<.001). We interpret this difference as indicating that L2 learners were erroneously led to an Instrument interpretation in sentences with Instrument prosody prior to reaching the end of the PP. Consequently, they may have experienced processing difficulty on hearing a non-instrument noun, and were led down a semantic garden-path.

Table 2. Analysis of looks to the instrument object for the duration of the final word

	Estimate	*Std. Error*	*t*-value	*p*-estimate
Intercept	-5.33	0.23	-23.08	<.001
Prosody	0.89	0.20	4.39	<.001
Noun Type	-0.64	0.20	-3.14	0.002
Trial	0.001	0.01	0.26	0.800
Prosody*Noun Type	0.01	0.41	0.03	0.974

3.2.2. Experiment 2 (uninformative filler prosody)

In Experiment 2, we investigated the extent to which L2 learners are sensitive to prosody-syntax mapping, and whether structural analysis in L2 processing is affected when prosodic boundary information does not always align with syntax. Previous studies exploring adaptation have shown that native speakers track how informative or reliable prosodic cues are in the experiment and adjust the extent to which they use prosody in structural analysis (Tanenhaus, Kurumada, & Brown, 2015; Roettger & Franke, 2017). The control group in the current study also showed that native speakers placed less weight on prosodic information in structural decision when the filler items in the experiment had a boundary tone that did not align with syntax (Nakamura et al., 2019). If L2 learners' knowledge about the likely mappings between linguistic information and incoming input is weaker compared to native speakers, L2 learners might be less sensitive to mismatch between prosody and syntax compared to native speakers. Thus, it is possible that L2 processing might be less disrupted when prosodic boundary information does not cue the syntactic structure of the sentence, thus "uninformative".

The same analyses used in Experiment 1 were conducted in Experiment 2. Fig. 5a shows the proportion of looks to the modifier object from the onset of sentence final word until the mean sentence offset. Table 3 summarizes the results from the model. The analysis of the proportion of looks to the modifier object (*the tiger with binoculars/popcorn* in Fig. 3) showed a main effect of Prosody ($p<.001$), so that participants looked at the modifier object significantly more with Modifier prosody than with Instrument prosody. In addition, there was a main effect of Trial ($p<0.05$), demonstrating that participants looked more at the modifier object as they experienced more trials in the experiment, regardless of prosody type. The results suggest that when the prosody in filler items was uninformative, L2 learners adopted the modifier interpretation more often than the instrument interpretation with Instrument prosody, as they were exposed to more uninformative prosody.

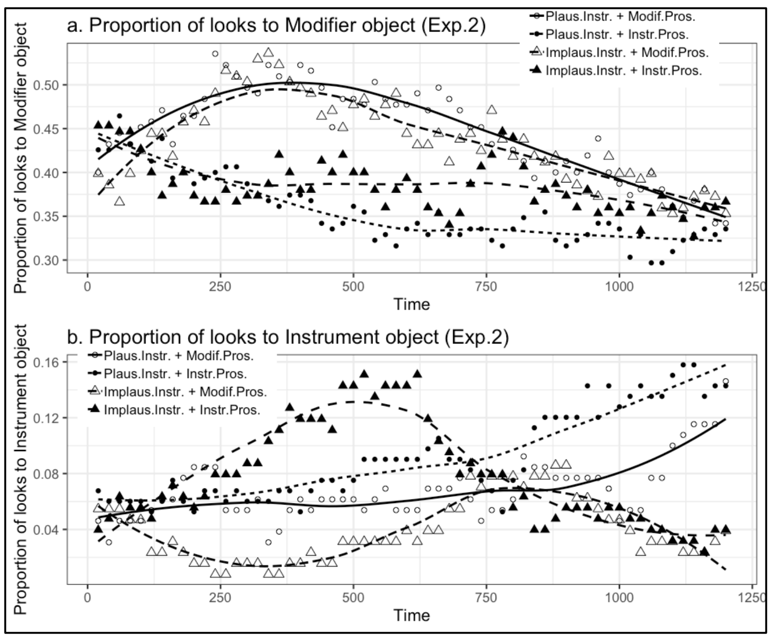


Fig. 5 Proportion of looks to (a) the modifier object and (b) the instrument object in Experiment 2

Table 3. Analysis of looks to the modifier object for the duration of the final word

	Estimate	*Std. Error*	*t*-value	*p*-estimate
Intercept	0.95	0.46	2.06	0.049
Prosody	-1.08	0.30	-3.60	<.001
Noun Type	-0.33	0.30	-1.10	0.270
Trial	0.59	0.18	3.28	<.01
Prosody*Instrument	0.71	0.60	1.18	0.237

We next analyzed the looks made to the instrument object. Fig. 5b shows the probability of looks to the instrument object (the instrument binoculars in Fig. 3) for the duration of the sentence final word. As in Experiment 1, we also asked whether there was a difference between conditions with different boundary locations in the Implausible instrument condition. Table 4 summarizes the results from the best-fitting model. The analysis again showed a main effect of Prosody ($p<0.01$); there were more looks to the instrument object with Instrument prosody than with Modifier prosody. The difference between the two types of prosody in the Implausible instrument conditions was not significant ($p=0.120$). The lack of statistical difference in the Implausible instrument conditions may suggest that participants in Experiment 2 recovered faster compared to participant in Experiment 1 after being misled to the instrument analysis in the Instrument prosody + Implausible instrument condition.

Table 4. Analysis of looks to the instrument object for the duration of the final word

	Estimate	*Std. Error*	*t*-value	*p*-estimate
Intercept	-4.95	0.29	-16.85	<.001
Prosody	0.63	0.24	2.61	<.01
Noun Type	-0.24	0.24	-0.99	0.325
Trial	0.02	0.15	0.16	0.878
Prosody*Noun Type	-0.14	0.48	-0.29	0.771

3.2.3. Combined analysis between Experiment 1 and 2

In order to compare the results of Experiments 1 and 2, we conducted a combined analysis. We first analyzed the looks made to the modifier object for the duration of sentence final word. The analysis showed an interaction between Experiment and Trial Order ($p<.05$), such that the Trial Order effect was significant only in Experiment 2. The results most likely suggest that when filler items had uninformative prosody, participants adopted the modifier analysis more regardless of prosody type as they heard more tokens in the experiment. L2 learners may have simply learned to ignore prosodic boundary information and focused instead on the picture that was compatible with both a modifier and an instrument interpretation. Alternatively, as the modifier object always received more looks than the instrument object, irrespective of the conditions, it is also possible that L2 learners adopted the modifier interpretation more when filler items had uninformative prosody because the modifier interpretation was somehow visually more salient than the instrument interpretation.

Another combined analysis on the looks to the instrument object was conducted in order to see if there was a difference in the size of garden-path effect in the Instrument prosody + Implausible instrument condition between the two experiments. With the native speaker control group, we observed an interaction between Prosody and Experiment in the Implausible instrument condition; the processing penalty arising from a mismatch between prosody and lexical input was smaller in Experiment 2 where filler items had uninformative prosody than in Experiment 1. However, the same combined analysis with our L2 learners showed no interaction between Prosody and Experiment. The results most likely suggest that, unlike native speakers who showed smaller garden-path effects in Experiment 2, the influence of uninformative prosody was weak with L2 learners.

4. General Discussion

Our results showed that L2 learners do indeed use prosodic boundary information to guide early parsing decisions, but at a delayed time course compared to native speakers. The results also suggest that L2 learners were sensitive to the prosody-syntax alignment and used prosodic information less in predictive structural analysis when the prosody for filler items was uninformative, as tested in Experiment 2. However, we found no evidence for the difference in the size of processing difficulty due to the mismatch between prosody and lexical compatibility with the preferred structure between experiments. Compared to native speakers, the extent to which L2 learners use prosodic information is impacted less by the overall consistency of other prosodic information in the experiment.

The results of the current study suggest that L2 learners make expectations in incremental structural building using prosody, which provide evidence against the view that L2 learners are incapable of making predictions during comprehension (e.g., Grüter & Rohde, 2013). Instead, our results suggest that although phonological information is used to anticipate upcoming structure across L1 and L2 populations (consistent with Ito, Pickering, & Corley, 2018), the time course at which L2 speakers use prosodic information may be delayed compared to native speakers. One possible cause for the delay could be the increased cognitive demand required when different sources of information must be integrated in order to anticipate upcoming structure. Indeed, many sentence processing models that incorporate a predictive component assume that parsing mechanism requires integrations of the constraints imposed by multiple sources of information such as syntax, semantic information, context, and prosody (e.g., Hale, 2006; Levy, 2008; McDonald, 2013). Due to the greater attentional demand for basic processing in L2, it is possible that L2 learners have reduced access to processing resources needed to make predictions about upcoming information, or are less able to rapidly adjust their expectations to the speaker's communicative intent. The results are also consistent with the claim that expectations for upcoming information are made based on general statistical regularities. Some accounts that explain language adaptation and learning propose that language users cope with linguistic variability by updating distributional statistics based on the exposure they receive (e.g., Farmer, Brown, & Tanenhaus, 2013; Chang, Dell, & Bock, 2006). It is possible that L2 learners may have limited access to distributional regularities due to the less overall exposure to linguistic input or to a weaker grammatical knowledge.

To conclude, our results provide evidence that prosodic boundary information is used in online structural building in L2 processing. L2 learners generated expectations based on prosody and visual information, and experienced processing difficulty when they heard information that was different from what they had anticipated. Our results further show that L2 learners are sensitive to prosody-syntax mapping, and use prosodic information less in structural analysis when prosodic boundary information does not align with syntax. It seems that L2 learners keep track of speaker-specific factors, such as how informative the speaker's prosodic cues are, and adjust the degree to which they use prosody to guide structure building processes.

References

Baayen, R. Harald, Davidson, Douglas J., & Bates, Douglas M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

Barr, Dale J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457-474.

Beckman, Mary & Pierrehumbert, Janet (1986). Intonational Structure in Japanese and English. *Phonology Yearbook*, *3*, 255-309.

Chang, Franklin, Dell, Gary S., & Bock, Kathryn (2006). Becoming syntactic. *Psychological Review*, *113*, 234–272.

Dahan, Delphine, Tanenhaus, Michael K., & Chambers, Craig G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *27*, 292-314.

De Jong, Nivja H. & Wempe, Ton (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior research methods*, *41*, 385 - 390.

Farmer, Thomas A., Brown, Meredith, & Tanenhaus, Michael, K. (2013). Prediction, explanation, and the role of generative models in language processing. *Behavioral and Brain Sciences*, *36*, 211–212.

Fine Alex B., Jaeger T. Florian, Farmer Thomas A., & Qian, Ting. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE 8*, e77661.

Grüter, Theres & Rohde, Hannah (2013). L2 processing is affected by RAGE: Evidence from reference resolution. Paper presented at the *12th conference on Generative Approaches to Second Language Acquisition (GASLA)*. University of Florida, FL.

Hale, John (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, *30*, 609–642.

Hwang, Hyekyung & Schafer, Amy J. (2006). Prosodic effects in parsing early vs. late closure sentences by second language learners and native speakers. In Rüdiger Hoffmann & Hansöjrg Mixdorff (Eds.), *Proceedings of Speech Prosody* 2006 (pp. 585-588). Dresden, Germany.

Ito, Aine, Pickering Martin J., & Corley, Martin (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1-11.

Ito, Kiwako, & Speer, Shari R. (2008). Anticipatory effect of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*, 541-573.

Jun Sun-Ah (2005). *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Levy, Roger (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126–1177.
MacDonald, Maryellen C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4, 226.
Martin, Clara D., Thierry, Guillaume, Kuipers, Jan-Rouke, Boutonnet, Bastien, Foucart, Alice, & Costa, Albert (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69, 574–588.
Nakamura, Chie, Arai, Manabu, & Mazuka, Reiko (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition*, 125, 317-323.
Nakamura, Chie, Harris, Jesse A., & Jun, Sun-Ah (2019). Anticipatory effect and the reliability of prosody in processing globally ambiguous sentences. Manuscript in preparation. University of California, Los Angeles and Japan Society for the Promotion of Science.
Nespor, Marina & Vogel, Irene (1986). *Prosodic Phonology*. Dordrecht: Foris.
Nickels, Stefanie & Steinhauer, Karsten (2016). Prosody-Syntax Integration in a Second Language: Contrasting Event-Related Potentials from German and Chinese Learners of English Using Linear Mixed Effect Models. Second Language Research, 34, 9-37.
Norris, Dennis & McQueen, James M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological Review*, 115, 357–395.
Roettger, Timo & Franke, Michael (2017). Evidential strength of intonational cues and rational adaptation to (un-)reliable intonation. Unpublished manuscript.
Schafer, Amy J., Speer, Shari R., Warren, Paul, & White, S. David (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29, 169-182.
Selkirk, Elizabeth (2000). The interaction of constraints on prosodic phrasing. In Merle Horne (Ed.), *Prosody: Theory and experiments* (pp. 231-262). Dordrecht: Kluwer.
Selkirk, Elizabeth (2007). Focus prominence drives focus phrasing: Bengali intonation revisited. In Chung-min Lee, Matthew Gordon and Daniel Büring (Eds.) *Topic and focus: Cross-linguistic perspectives on Intonation and Meaning* (pp. 217-246). Dordrecht: Springer.
Selkirk Elizabeth (2011). The syntax-phonology interface. In John Goldsmith, Jason Riggle, and Alan Yu (Eds.) *The Handbook of Phonological Theory*, 2nd edition (pp. 435-84). Oxford: Blackwell Publishing.
Shattuck-Hufnagel Stephanie & Turk, Alice E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*, 193-247.
Snedeker, Jesse, & Trueswell, John (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*, 103-130.
Tanenhaus, Michael K., Kurumada, Chigusa, & Brown, Meredith (2015). Prosody and intention recognition. In Lyn Frazier and Edward Gibson (Eds.) *Explicit and Implicit Prosody in Sentence Processing: Studies in Honor of Janet Dean Fodor*, (pp. 99-118). Studies in Theoretical Psycholinguistics, 46. Berlin: Springer.
Truckenbrodt, Hubert (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry*, *30*, 219–255.
Weber, Andrea, Grice, Martine, & Crocker, Matthew W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, *99*, 63-72.