BUCLD 45 Proceedings
To be published in 2021 by Cascadilla Press
Rights forms signed by all authors

Adaptation to Atypical Contrastive Accent: The L2 Advantage

Chie Nakamura, Jesse A. Harris, and Sun-Ah Jun

1. Introduction

A growing number of visual world studies have reported that listeners process prosodic cues rapidly enough to predict upcoming speech, directing their eyes to a speaker-intended referent before the onset of the referring noun (Ito & Speer, 2008; Snedeker & Trueswell, 2003; Weber, Braun, & Crocker, 2006). These studies have converged on the idea that prosodic cues are used to guide processing decisions very early stages, and further enter into a larger debate regarding the prediction of words and structures in real time comprehension. The general conclusion is that language users draw on general statistical knowledge of how prosody functions in conjunction with other sources of information, such as contextual and world knowledge, to generate predictions about what information is likely to come next. The specific influence of prosody on predictive inferences – often referred as the *anticipatory use of prosody*, has been observed in multiple levels of representations, including structural prediction (e.g., Nakamura, Arai, & Mazuka, 2012; Nakamura, Harris, & Jun, 2019; Weber, Grice, & Crocker, 2006) and the interpretation of discourse and information structure (e.g., Ito & Speer, 2008; Dahan, Tanenhaus, & Chambers, 2002; Kurumada, Brown, Bibyk, Pontillo, & Tanenahus, 2014).

More recently, a number of studies have manipulated whether or not a particular speaker's prosody follows conventional use (e.g., boundaries or pitch accents that realize typical relationships between prosody and structure or meaning) and the extent that prosodic marking is informative to the listener. The results suggest that listeners monitor how specific speakers use prosodic information, and adjust the extent to which they rely on this information in anticipatory processing (Nakamura, Harris, & Jun, 2019; Roettger & Franke, 2019; Roettger & Rimland, 2020; Tanenhaus, Kurumada, & Brown, 2015). The observation that comprehension strategies change in response to linguistic input is sometimes taken as evidence for *language adaptation*, in which language users

---

* Chie Nakamura, Waseda University, Tokyo, Japan cnakamura@aoni.waseda.jp; Jesse A. Harris, University of California, Los Angeles; Sun-Ah Jun, University of California, Los Angeles. We thank the members of Language Processing Lab at UCLA and BLIT lab at Waseda University for their assistance in data collection. Portions of the research were previously presented at Speech Prosody 2020, and we thank the reviewers and the audience for their comments and questions. This work was supported by JSPS KAKENHI Grant Number JP19H01279.

track statistical properties of the speech situation to dynamically update their language processing strategies, allowing them to better extract the intended message from a noisy speech signal.

Compared to L1 processing, the role of anticipation and adaptation in second language (L2) processing is much less well understood. Despite its intuitive advantages, anticipatory processing might also place additional cognitive demands on the language processing, e.g., by increasing the sources of information that guide early stages of processing. An increased cognitive demand for anticipatory processing raises the possibility that predictive strategies might be avoided in populations thought to have limited access to cognitive reserves, such as older adults (Federmeier, 2007; Federmeier, Kutas, & Schul, 2010) or non-native speakers (Kaan, 2014; Lew-Williams & Fernald, 2010; Martin, Thierry, Kuipers, Boutonnet, Foucart, & Costa, 2013; Grüter, Rohde, & Schafer, 2014). For instance, the Reduced Ability to Generate Expectations (RAGE) hypothesis, proposed by Grüter, Rohde, & Schafer (2014), argues that L2 learners are so overtaxed by integrative processes that they have little processing resources remain to predict upcoming information compared to native speakers.

However, a number of recent eye-tracking studies investigating L2 processing report that some L2 learner populations are able to utilize prosodic cues in anticipatory processing (Nakamura, Harris, Jun, & Hirose, 2019; Perdomo & Kaan, 2019). These findings provide suggestive evidence against the view that L2 learners are categorically unable to engage in anticipatory processing (Lew-Williams & Fernald, 2010; Mitsugi & MacWhinney, 2016), and are more consistent with the view that L2 anticipatory processing is graded rather than an all or none process (Kaan, 2014; Curcic, Andringa, & Kuiken, 2018; Ito & Pickering, 2021).

One example comes from Nakamura et al. (2019), who compared how native speakers and L2 learners of English used prosodic boundary information when processing a structurally ambiguous sentence in a visual scene. Listeners were presented with sentences like (1), in which the PP *with the binoculars* is structurally ambiguous between a modifier (1a) and an instrument (1b) interpretation. The placement of a prosodic boundary (%) can serve to disambiguate the structure by aligning with the syntactic constituency of the sentence (Selkirk, 2000, 2007, 2011). They reasoned that, if prosodic boundary placement disambiguated the sentence structure (i.e., a prosodic boundary cued by a rising tone, L-H%, and phrase-final lengthening; (2)), then mismatches between the syntactic and prosodic structure in (2b) would elicit a processing cost. They created semantic garden path effects by manipulating whether the final noun was plausible (*binoculars*) or implausible (*popcorn*) as an instrument.

(1) The boy will see the tiger with the binoculars / popcorn.
  a. *Modifier paraphrase:*
  The boy will see the tiger, who has the binoculars / popcorn.
  b. *Instrument paraphrase:*

The boy will use the binoculars / # popcorn to see the tiger.

(2) a. *Modifier prosody:*

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

b. *Instrument prosody:*

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

The eye-movement patterns revealed that implausible instruments (*popcorn*) elicited a processing cost when paired with instrument prosody (2b) for both native speakers and L2 learners of English. Although the effect in L2 learners was delayed relative to native speakers, the result indicates that L2 learners are sensitive to the placement of prosodic boundaries in anticipatory language processing.

Nakamura et al. (2019) presented the same experimental materials embedded into filler items containing atypical or unreliable prosodic cues, in which prosodic boundary information did not meaningfully align with syntactic constituencies, to a separate group of native speakers and L2 learners. For example, a boundary cued by a low tone, L-L%, was placed between a determiner and a noun, breaking the correspondence between prosody and syntax, as shown in (3). Both L1 and L2 groups quickly adjusted their expectations on critical trials, appearing to rely less on prosodic information as the experiment progressed. However, the effect of uninformative prosody was weaker for L2 learners compared to native speakers.

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

The reduced negative impact of atypical prosody on L2 processing might be explained by a weaker mapping between prosody and syntax for L2 learners compared to native speakers. As proposed in some models of L1 language processing, language users are able to cope with linguistic variability by updating distributional statistics based on the exposure they receive (Farmer, Brown, & Tanenhaus, 2013; Chang, Dell, & Bock, 2006). As adult L2 learners are exposed to far less linguistic input, and at a later stage in learning, they may have relatively loose associations between prosody and syntax. In terms of Nakamura et al.'s (2019) study, both groups of participants appeared to have down weighted the role of prosodic boundary information when prosodic information was atypical, albeit to different degrees. If native speakers initially had a strong expectation for the syntax and prosody to align, revising this expectation would result in drastically different processing behavior. In contrast, a weaker expectation or mapping for L2 learners would explain their reduced response to atypical prosody.

This difference raises the interesting possibility that a weaker mapping between prosody and its conventional use might allow L2 learners to adapt more quickly to a speaker-specific use of prosody when the prosodic marking is unconventional or atypical. As suggested by previous studies, there is a general consensus that L2 learners have a weaker mapping between prosody and meaning compared to native speakers (e.g., Akker & Cutler, 2003; Fivela 2012; Dolan, 2015; Nakamura et al., 2019). If L2 learners have less rigid or less defined prosodic mappings, they could be better positioned to accept atypical uses of prosody as a new category compared to native speakers, and may even actively use atypical prosody to anticipate upcoming information.

The present study addresses the following two issues: first, whether L2 learners would use prosodic information in anticipatory processing as in previous studies, and, second, whether L2 learners would better learn experiment-specific uses of atypical prosody to correctly predict the upcoming object. We compared the eye movements of native English speakers and L2 learners of English whose L1 is Japanese. In two visual-world eye-tracking experiments, we first tested whether a conventional / typical use of contrastive accent ((4); *red cat* followed by *PURPLE cat*) increased anticipatory looks to the target object (i.e., fixations on a to-be-mentioned object before the onset of the object word). We then tested how listeners respond to an unconventional / atypical use of contrastive accent ((5); *red cat* followed by *PURPLE pig*) in Experiment 2. The same set of visual arrays were displayed in each experiment, e.g., Fig. 1. In the following sections, we first report the two experiments in Section 2 and 3, respectively. We then discuss the results of the two experiments in Section 4.

(4) *Conventional use of contrastive accent (Experiment 1)*
First, find the red cat.
Next, find the PURPLE$_{L+H*}$ cat.

(5) *Atypical use of contrastive accent (Experiments 2)*
First, find the red cat.
Next, find the PURPLE $_{L+H*}$ pig.

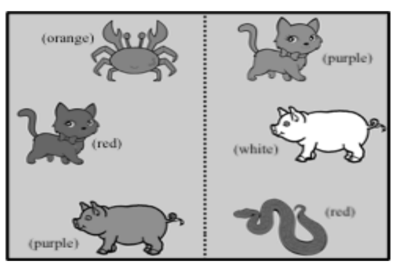

Fig. 1. Visual array presented with (4) and (5). The color in parenthesis indicates the color used for each object

2. Experiment 1

In Experiment 1, we tested whether the conventional use of contrastive accent (L+H*) on the adjective *purple* (6b) leads to anticipatory looks to a target object (e.g., a purple cat) as compared to the control condition presented with new information accent (H*) (6a).

(6) First, find the red cat.
  a. Next, find the purple$_{H*}$ cat.
  b. Next, find the PURPLE$_{L+H*}$ cat.

Stimuli

Thirty-six experimental items were created. Each item consisted of a sound file of a sentence and a corresponding visual scene, prepared with clip art images, as in Fig. 1. The auditory stimuli were recorded by a male native speaker of English trained in English ToBI (Beckman, Hirschberg, & Shattuck-Hufnagel, 2005). Fig. 2 shows the f0 contours of the sentence (6) in each condition. The position of the objects was counter-balanced across the items. The 36 experimental items were pseudo-randomly interspersed with 54 unrelated filler items in two counterbalanced lists.

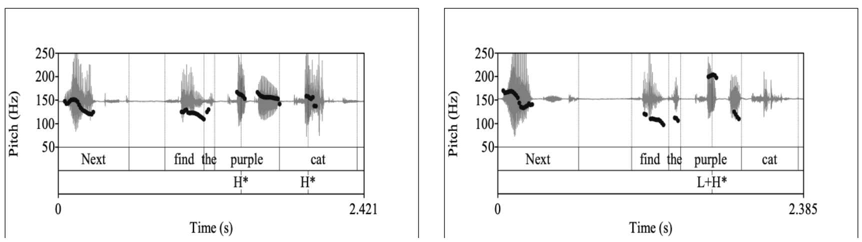

Fig. 2. Waveform and pitch track showing the accent type for (4a, left) and (4b, right)

Procedure

Participants were instructed to listen to the sentences carefully while attending to the picture on the computer monitor. As soon as the sentence ended, participants were instructed to indicate the position of the second-mentioned object using the left and right bumpers on a USB gamepad. All participants performed above 98% on the post-sentence identification task in all experiments reported in this paper. The visual array was presented for 3000ms prior to the onset of the sentence to allow listeners to familiarize themselves with the objects. Eye movements were recorded with EyeLink 1000 Plus (SR Research) at a sampling rate of 500 Hz. A 5-point calibration was conducted at the beginning of the experiment and as needed. Drift correction was performed before each trial. Experimental sessions lasted approximately 30 minutes.

Data Analysis

For the eye-movement analysis, we summed the number of gazes to each entity in the scene and calculated the logit of looks to each entity out of looks to all the objects in the scene, including the background, for the period of interest. Statistical analyses for the duration of a predetermined *anticipatory time window* (from the onset of the color adjective until the minimum onset of the target noun, 400ms) were conducted using linear mixed-effect regression models (Baayen, Davidson, & Bates, 2008). Accent Type (H* or L+H*) was included as a fixed effect predictor in all models. Additional models were constructed to investigate whether participants' performance changed over the course of the experiment. In these models, the 90 trials were divided into 3 blocks and included as an additional interactive factor in the model (Block). By-participant and by-item effects were modeled as random intercepts, after random slope models were found to overfit the data. The best fitting model for each analysis was computed by eliminating non-significant predictors from the model using a backward selection approach.

2.1. Native speaker results (Experiment 1)

We collected data from thirty-six native speakers of English from the student community at the University of California, Los Angeles as a control group. Table 1 summarizes the results from the best-fitting model in the anticipatory time window. Fig. 3 (left) shows L1 participants' proportion of looks to the target object from the onset of the color adjective. The analysis on the looks made to the target object showed a main effect of Accent Type ($\beta = 2.37$, $SE = 1.05$, $t = 2.25$, $p < .05$); more looks to the target object were observed with contrastive accent than with the control, new information accent. Crucially, the effect was observed in the anticipatory time window, before the onset of the following noun. No effect of Block was observed. As predicted, native speakers used contrastive accent to program an eye-movement in anticipation of the upcoming word that was most likely to contrast with the previous word in the visual array, supporting previous findings (Weber et al., 2006; Ito & Speer, 2008).

2.2. L2 learner results (Experiment 1)

Thirty Japanese L1 speakers, learning English as an L2, participated in the experiment as the L2 group. Participants were recruited from the student community at Waseda University. 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). Fig. 3 (right) shows L2 participants' proportion of looks to the target object from the onset of the color adjective. Consistent with the results of the native speakers, there was a main effect of Accent Type for the duration of the anticipatory time window ($\beta = 1.35$, $SE = 0.65$, $t = 2.26$, $p < .05$). This effect indicates that L2 learners made more anticipatory looks to the target object when adjectives bore contrastive accent compared to new information accent.

An alternative explanation is that L2 learners looked to the object corresponding to the accented word simply because the adjective was phonetically more prominent. This possibility would predict that the competitor object (e.g., purple pig) would attract as many looks when adjective bore contrastive accent. However, an analysis of looks to the competitor object revealed no effect of Accent Type ($p = 0.18$) in the anticipatory time window. Thus, both L2 learners and native speakers alike appeared to anticipate the likely referent of the upcoming noun on the basis of the type of accent on the adjective.

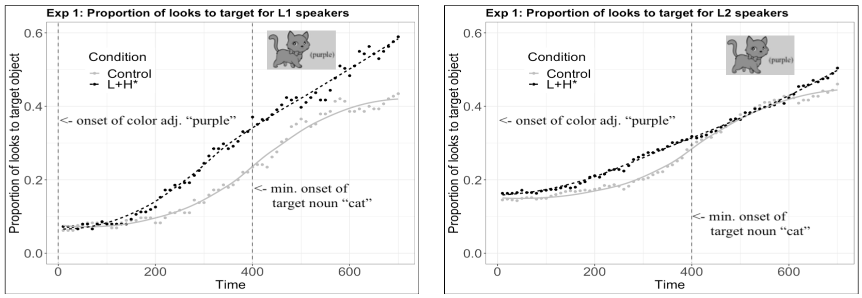

Fig. 3. Proportion of looks to the target object from the onset of the color adjective to 700ms in Exp. 1 for Native speakers (left) and L2 learners (right), time-locked to the onset of color adjective (e.g., "purple") of each item.

We observed a marginally significant effect of Block for L2 learners, resulting in increased anticipatory looks to the target object as the experiment progressed, regardless of prosodic type ($\beta = 0.49$, $SE = 0.26$, $t = 1.90$, $p = 0.06$). There was also a marginally significant interaction of Accent Type and Block ($\beta = -0.55$, $SE = 0.30$, $t = -1.83$, $p = 0.07$). As shown in Fig. 4 (left), the direction of the interaction suggests that the effect of Accent Type was reduced in later trials. These results will be discussed in the analysis of the referent noun time window below.

Referent noun time window

The *referent noun time window* was defined as the onset of the referent noun until the minimum offset of the sentence (450ms). An analysis of looks to the target object in this window revealed an interaction between Block and Prosody ($\beta = 0.65$, $SE = 0.33$, $t = 1.98$, $p < .05$). L2 learners looked at the target object less with new information accent as the experiment proceeded (Fig. 4 right), suggesting that they continued to fine-tune their understanding of Accent Type as the experiment progressed.

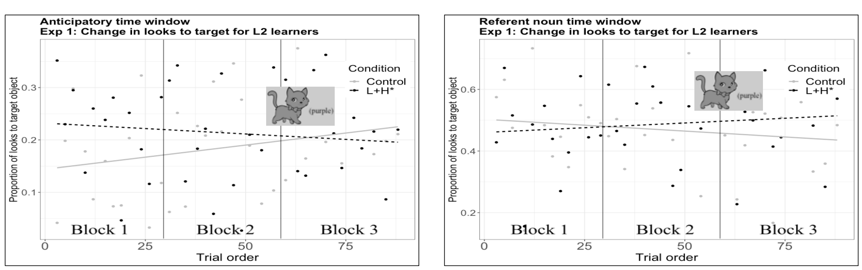

Fig. 4 Change in looks to the target object in the anticipatory time window (left) and the referent noun time window (right) across trials

3. Experiment 2

Several recent studies have explored whether listeners track how informative or reliable prosodic cues are, and, if so, whether they adjust the extent to which prosody is used to disambiguate sentences during online comprehension (e.g., Tanenhaus, Kurumada, & Brown, 2015; Roettger & Franke, 2019; Nakamura, Harris, & Jun, submitted). However, it remains unclear how far listeners will go to adjust to speaker-specific, unconventional use of prosody. A conservative possibility is that listeners simply reduce the extent to which they engage in anticipatory processing, so that they rely more on lexically-driven strategies that favor bottom-up input. Another possibility is that they learn atypical prosodic marking patterns that are associated with a particular speaker and begin to make anticipatory judgments on the basis of the learned category.

In order to address this issue, Experiment 2 tested how listeners respond to an atypical use of contrastive accent (7b) compared to a control condition with new information (H*) accent (7a). The same visual scenes as in Experiment 1 were used in Experiment 2.

(7) First, find the red cat.
  a. Next, find the purple$_{H*}$ pig.
  b. Next, find the PURPLE$_{L+H*}$ pig.

If listeners adjust their use of prosodic cues based on how informative the prosody is, listeners should place less weight on the kind of accent that appears on the adjective in this experiment. This scenario predicts that both native speakers of English and L2 learners would make fewer anticipatory eye-movements in Experiment 2 compared to the participants in Experiment 1. If, however, L2 learners have weaker prosodic mappings or more malleable representations of prosodic categories in their L2 than native speakers do, then L2 learners should show greater adaptation to atypical, unconventional uses of contrastive accent compared to native speakers. In this case, L2 learners would learn the speaker-specific atypical use of contrastive accent in the experiment and eventually predict the correct object with the atypical use of contrastive accent in Experiment 2.

3.1. Native speaker results (Experiment 2)

Thirty-five native speakers of English from the same population as in Experiment 1 participated in Experiment 2. Fig. 5 (left) shows the proportion of looks that native listeners made to the target object from the onset of the color adjective. The results of the analysis on the looks made to the target object (e.g., purple pig) in the anticipatory time window showed no effect of Accent Type ($p = 0.52$) or Block ($p = 0.24$). Analyses of looks to the competitor object (e.g., purple cat) showed no effect of Accent Type ($p = 0.17$) or Block ($p = 0.80$). The results suggest that native speakers did not make anticipatory eye-movements to the target object nor to the competitor object when presented with atypical uses of contrastive accent.

3.2. L2 learner results (Experiment 2)

Twenty-nine L2 learners of English from the same L1 Japanese-speaking population as in Experiment 1 participated in the experiment. The analyses were conducted as before. Fig. 5 (right) shows the proportion of looks to the target object from the onset of the color adjective. There was a marginal main effect of Accent Type in the anticipatory time window ($\beta = 0.44$, $SE = 0.24$, $t = 1.83$, $p = 0.07$); L2 learners tended to look at the target object (purple pig) more often in the L+H* condition than in the H* condition. As discussed below, the trend possibly reflects a greater sensitivity to fine-grained phonetic cues for L2 learners (Lorch & Meara, 1989).

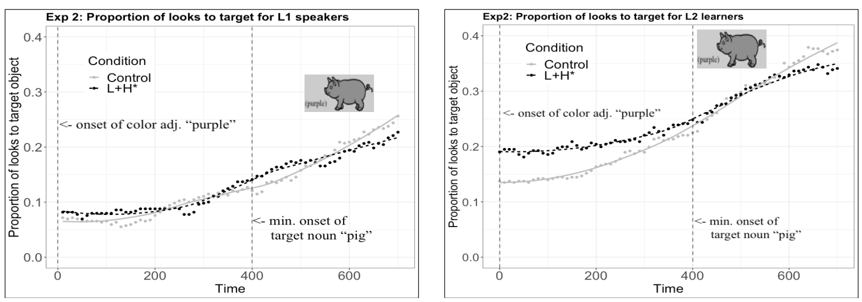

Fig. 5. Proportion of looks to the target object from the onset of the color adjective to 700ms in Exp. 2 for native speakers (left) and L2 learners (right), time-locked to the onset of color adjective (e.g., “purple”) of each item.

The results also showed a main effect of Block ($\beta = 0.70$, $SE = 0.18$, $t = 3.90$, $p < .001$), such that L2 learners looked at the target object more often as the experiment progressed, regardless of Accent type, as illustrated in Fig.6. The results suggest that, with atypical use of contrastive accent, L2 learners adapted their processing strategy as they experienced more trials.

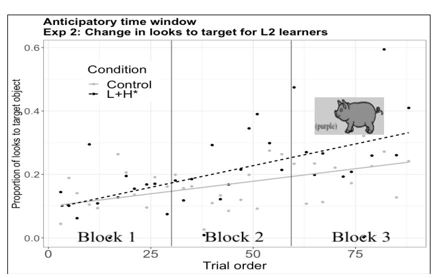

Fig. 6. L2 learners’ change in looks to the target object in the anticipatory window across trials in Exp. 2.

Pre-adjective time window

The plots for native speakers and L2 learners from the main analysis in Fig. 5 suggest that L2 learners were already fixating to the correct target object with atypical use of contrastive accent before the onset of the color object. This somewhat unexpected pattern raises another interesting possibility: L2 learners might have used additional acoustic cues that are not typically used by native speakers in anticipatory processing. In English intonation, L+H* and H* are known to differ phonetically by the f0 rising slope (i.e., sharp vs. smooth rise), but they also differ in the f0 on the syllable before the f0 peak (i.e., lower f0 in L+H*) (Beckman et al., 2005). Thus, it is possible that the phonetic difference between the two conditions before the peak influenced L2 learners' anticipatory eye-movements. To further explore this possibility, we first conducted an acoustic analysis to compare the f0 for the duration of "find the" segment in "Next, find the purple pig" in items used in Experiment 2. The results of the acoustic analysis showed that the L+H* condition had lower f0 before the peak (mean pitch peak = 95Hz, $SD = 12$), $t(35) = 6.44$, $p < .001$) compared to the H* condition (mean pitch peak = 109Hz, $SD = 4$). Next, we analyzed the looks made to the target object in the *pre-adjective time window* (from the onset of "find the" in each item until the minimum onset of the color adjective, 400ms) using the same model used in the main analysis. There was a marginal effect of Accent Type ($p = .06$) and a main effect of Block ($p < .05$). The marginal effect of Accent Type might reflect a greater sensitivity to fine-grained phonetic cues for L2 learners. The main effect of Block is compatible with the claim that L2 learners are sometimes able to adopt different processing strategies than native speakers.

4. General Discussion

The current study addressed two central questions in anticipatory language processing. The first question concerned whether L2 listeners use contrastive pitch accent to anticipate an upcoming referent. Previous research has indicated that L1 listeners make rapid inferences in speech processing on the basis of prosodic contour, and anticipate the referent that will be mentioned next in a visual scene (Ito & Speer, 2008; Weber et al., 2006). After confirming that L1 listeners use contrastive accent to anticipate upcoming referents with our materials and task, we found that L2 listeners did so as well. This finding was interpreted as evidence against the strong version of the RAGE hypothesis (Grüter, et al., 2014; see also Lew-Williams & Fernald, 2010; Mitsugi & MacWhinney, 2016 for related accounts). Our results are consistent with previous studies reporting an anticipatory use of prosody in L2 processing, providing broad support for the view that elements of L2 processing are proactive (Hopp, 2013; Hopp, 2016; Kaan, 2014; Foucart, Martin, Moreno, & Costa, 2014).

The second, and related, question explored whether L1 and L2 listeners adjust to atypical uses of contrastive accent and adapt their anticipatory processing strategies. On the one hand, L1 listeners have vastly more experience with the target language than L2 listeners, exposing themselves to considerable variation among speakers. This exposure may allow them to quickly create a speaker-specific model, increasing the chance of adaptation (e.g., Norris, McQueen, & Cutler, 2015). Without such experience, L2 listeners may not have strong expectations about the conventional use of prosodic information, making adaptation much more difficult. On the other hand, assuming that L2 listeners have weaker prosodic mappings or weaker categorical boundaries for prosodic categories in their L2 (e.g., Akker & Cutler, 2003; Fivela 2012; Dolan 2015; Nakamura et al., 2019), L2 listeners might instead be better positioned to adapt to speaker-specific variation, compared to L1 listeners.

Our Experiment 2 found that L2 listeners showed evidence of adaptation, whereas L1 listeners did not. Even though the use of contrastive accent was not typical in this experiment, L2 listeners looked more to the referent during the adjective. Together with the change in looks over the course of the experiment, it appears that L2 listeners learned the atypical use of contrastive accent, potentially adjusting their prosodic categories. Some language processing accounts suggest that language users have distributional regularities of how linguistic information should be used (e.g., Farmer et al., 2013; Chang et al., 2006). It is possible that L2 learners have limited access to distributional information on conventional prosodic categories, and a looser mapping between prosody and meaning allowed them to quickly learn the experiment-specific, atypical uses of contrastive accent as a new category and to actively use the information in anticipatory processing.

However, we wish to briefly consider two alternative conclusions here. First, there is considerable variation among speakers regarding the tonal production and perception of H* and L+H* accents in English (Ladd, 2008; Ladd & Morton, 1997; Ladd & Schepman, 2003), and phoneticians are divided as to the categorical or gradient nature of these accent types. Thus, it is possible that our L1 listeners in Experiment 2 happened to not distinguish between the two accents. This possibility, however, seems unlikely given the results from Experiment 1, which drew subjects from the same population. A second, and more plausible, alternative is that L1 listeners quickly inferred that the use of L+H* was unreliable and was not informative, and immediately began treating it on par with H* accent. It is possible that L1 listeners determined that L+H* was not a meaningful tonal expression for the speaker within the first few trials; in that case, there would have been no evidence of learning throughout the study.

As discussed, there is already a wealth of evidence that L1 listeners adapt to a range of phonological levels, ranging from phonetic (Clayards, Tanenhaus, Aslin, & Jacobs, 2008; Kleinschmidt & Jaeger, 2015; Kraljic & Samuel, 2008; Maye, Aslin, & Tanenhaus, 2008; Norris, et al., 2015) to prosodic (Kurumada et al., 2014) representations. Though not necessary for adaptation (Liu & Jaeger, 2018), L1 listeners appear to accommodate speaker-specific productions better if an explanation of the deviation is provided (e.g., Nakamura, Harris, & Jun, 2019). It is therefore possible that L1 listeners simply required more motivation or insight into the state of the speaker before initiating adaptation. Although we are unable, at this point, to determine which possibility is correct, it is clear that L1 and L2 listeners responded very differently to the use of atypical accent in Experiment 2, a finding that warrants further investigation.

Unexpectedly, L2 listeners showed a sensitivity to accent type prior to the adjective on which the accent was fully realized. We discussed the possibility that L2 learners may have used f0 differences to distinguish L+H* from H* prior to the full realization of pitch accent. The prosodic profile of the L1 may be relevant in determining what kinds of representations are available in an L2 through L1-transfer. The L1 of our subjects was Japanese, a lexical pitch accent language in which the location or type of accent can distinguish lexical meanings via f0 height (e.g., Kawahara, 2006, for review). For example, pitch accent on the initial syllable of *áme* (*rain*) is lexically distinct from the word *amé* (*candy*), which places an accent on the final syllable. As they must have been familiar with the importance of accent type and location in their L1, the L2 learners in our sample may have paid a great deal of attention to fine-grained phonetic details of the stimuli. Due to L1 transfer, our L2 participants may even have been more sensitive to f0 cues than native English speakers, who did not seem to use f0 height before the adjective as an early cue of contrastive accent. Further study is needed to assess this possibility, and L2 listeners from L1 languages besides Japanese may perform differently.

To conclude, we have provided evidence that L2 listeners do engage in anticipatory processing. Further, they may even exhibit selective advantages for prediction depending on what features are available from their L1. While processing an L2 is certainly challenging, it appears that L2 listeners may, at times, outperform L1 counterparts, under select conditions. And while additional experimentation is required to test the merits and limits of this proposal, the general view is compatible with an understanding of second language comprehension that invites continual interplay between L1 and L2 languages, resulting in a highly plastic perceptual system (e.g., Cutler, 2012).

References

Akker, E., and Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism, 6,* 81–96.

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

Beckman, M., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In: Jun, S.-A. (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 9-54), Oxford: Oxford University Press.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*, 234–72.

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition, 108,* 804-809.

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning, 69*, 42–71.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. MIT Press.

Dahan, D., Tanenhaus, M.K., & Chambers, C.G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*, 292–314.

Dolan, J. (2015). How second language English learners interpret sarcasm in English: A survey. *Schwa: Language and Linguistics, 13*, 11-26.

Federmeier, K.D. (2007), Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491-505.

Federmeier, K.D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115*, 149–161.

Farmer, T.A., Brown, M., Tanenhaus, M.K. (2013). Prediction, explanation, and the role of generative models in language processing. *Behavior and Brain Sciences, 36,* 211-212.

Fivela, B. G. (2012) Testing the perception of L2 intonation. In: Busà, M. G., & Stella, A. (Eds.) *Methodological Perspectives on Second Language Prosody*. (pp. 17-30). Padova: CLEUP.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1461–1469.

Grüter, T. & Rohde, H., & Schafer, J. A. (2014). The role of discourse-level expectation in non-native speaker's referential choices. In: Orman, W. & Valleau, M. J. (Eds.), *BUCLD 38 Proceedings* (pp. 179-191). Somerville, MA: Cascadilla Press.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*, 33–56.

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research, 32*, 277–307.

Ito, A., & Pickering, M. (2021). Automaticity and prediction in non-native language comprehension. In: Grüter, T., & Kaan E. (Eds.), *Prediction in Second-Language Processing and Learning*. John Benjamins.

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

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism, 4*, 257–282.

Kawahara, S. (2015). The phonology of Japanese accent. In: Kubonozo, H. (Ed.), *The Handbook of Japanese Phonetics and Phonology* (pp. 445-492), de Gruyter.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*, 148–203.

Kraljic, T., & Samuel, A. G. (2011). Perceptual learning evidence for contextually-specific representations. *Cognition, 121*, 459–465.

Kurumada, C., Brown, M., Bibyk, S., Pontillo D.F., Tanenhaus, M.K. (2014). Is it or isn't it: Listeners make rapid use of prosody to infer speaker meanings. *Cognition*, *133*, 335-342,

Ladd, D. R. (2008). *Intonational Phonology.* 2nd edition. Cambridge University Press.

Ladd, D. R., & Morton, R. (1997). The perception of intonational emphasis: continuous or categorical? *Journal of Phonetics, 25*, 313-342.

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics, 31*, 81-112.

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender- marked articles by native and non-native Spanish speakers. *Journal of Memory and Language, 63*, 447–464.
Lorch, M., & Meara, P. (1989). How people listen to languages they don't know. *Language Sciences, 11*, 343-353.
Martin, C.D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language, 69*, 574–588.
Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition, 19*, 19–35.
Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2012). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science, 32,* 543-562.
Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition*, *125*, 317-323.
Nakamura, C., Harris, J. A., & Jun, S.-A. (2019). Listeners' beliefs influence prosodic adaptation: Anticipatory use of contrastive accent during visual search. *Proceedings of the 19th International Congress of Phonetic Sciences*, 447-451.
Nakamura, C., Harris, J. A., & Jun, S.-A. (submitted). Integrating prosody in anticipatory language processing: How listeners adapt to unconventional prosodic cues. Ms. Waseda University and University of California, Los Angeles.
Nakamura, C., Harris, J. A., Jun, S.-A., & Hirose, Y. (2019). L2 adaptation to unreliable prosody during structural analysis: A visual world study. In: Brown, M. M. & Dailey, B. (Eds.), *BUCLD 43 Proceedings* (pp. 454-468). Somerville, MA: Cascadilla Press.
Norris, D., McQueen, J., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition, *Language, Cognition and Neuroscience, 31*, 4-18.
Perdomo, M., & Kaan, E. (2019). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research*, 1-27.
Roettger, T. & Franke, M. (2019). Evidential strength of intonational cues and rational adaptation to (un-)reliable intonation. *Cognitive Science*, e12745.
Roettger, T. & Rimland, K. (2019). Listeners' adaptation to unreliable intonation is speaker-specific. *Cognition, 204*, e104372.
Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*, 103-130.
Tanenhaus, M.K., Kurumada, C., & Brown, M. (2015). Prosody and intention recognition. In: Frazier, L. & Gibson, E. (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.
Weber, A., Braun, B., & Crocker, M.W. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and speech 49*, 367-392
Weber A., Grice M., & Crocker, M.W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition*, *99*, B63-72.