

Research Article

USING PROSODY TO PREDICT UPCOMING REFERENTS IN THE L1 AND THE L2 THE ROLE OF RECENT EXPOSURE

Anouschka Foltz  *

University of Graz

Abstract

While monolingual speakers can use contrastive pitch accents to predict upcoming referents, bilingual speakers do not always use this cue predictively in their L2. The current study examines the role of recent exposure for predictive processing in native German (L1) second language learners of English (L2). In Experiment 1, participants followed instructions to click on two successive objects, for example, *Click on the red carrot/duck. Click on the green/GREEN carrot* (where CAPS indicate a contrastive L + H* accent). Participants predicted a repeated noun following a L + H* accent in the L1, but not in the L2, where processing was delayed. Experiment 2 shows that after an exposure period with highly consistent prosodic cues, bilinguals engaged in predictive processing in both their L1 and L2. However, inconsistent prosodic cues showed different effects on bilinguals' L1 and L2 predictive processing. The results are discussed in terms of exposure-based and resource-deficit models of processing.

INTRODUCTION

Abundant evidence in the literature shows that listeners engage in predictive processing in their native language (Kamide, 2008). This suggests that listeners not only integrate incoming sentence material into the phrase they are currently processing but also use the

 Data collection for this study was supported by the Collaborative Research Centre 673 “Alignment in Communication” and the Cognitive Interaction Technology—Center of Excellence (CITEC) at Bielefeld University and by the Department of Languages, Literatures & Linguistics at Bangor University. I additionally acknowledge the financial support by the University of Graz. I would like to thank Prof. Katharina Rohlffing for the use of her lab, and Luci Fillinger for her diligence in collecting the data. I would also like to thank Prof. Chong Gu for his help and advice with the statistical analyses and two anonymous reviewers for their careful review and constructive comments on the article. All remaining errors are, of course, my own.

* Correspondence concerning this article should be addressed to Anouschka Foltz, Institute of English Studies, University of Graz, Heinrichstraße 36/II, 8010 Graz, Austria. E-mail: anouschka.foltz@uni-graz.at

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

already available information to make predictions as to which words, syntactic structures, and so forth come next. Native listeners can use various different linguistic cues and real-world information to make these predictions (Boland, 2005; Kamide et al., 2003; Lau et al., 2006; Lew-Williams & Fernald, 2010; Weber et al., 2006b) and can predict information at various levels of linguistic representation (Boland, 2005; DeLong et al., 2005; Lau et al., 2006; Van Berkum et al., 2005).

There is also abundant evidence that nonnative (L2) speakers do not engage in predictive processing to the same extent as native (L1) speakers (Kaan, 2014). This general finding is reflected in Grüter et al.'s (2017) hypothesis that L2 learners have Reduced Ability to Generate Expectations (RAGE). However, it is not yet clear in which situations and why L2 learners do or do not engage in predictive processing. Specifically, L2 learners may differ from native speakers in their use of predictive processing in terms of the cues that they use for prediction and/or the levels of representation that they predict. Individual learner differences may also contribute to whether or not learners engage in predictive processing. The current study focuses on the use of prosody, specifically contrastive pitch accents, as a cue to predict upcoming referents in German–English bilinguals' L1 and L2. Furthermore, the current study explores the effect of recent exposure to inconsistent versus consistent prosodic cues on bilinguals' ability to use prosody as a cue to predict upcoming referents.

USING PROSODY FOR PREDICTION IN THE L1

Studies with adult native speakers of various languages suggest that native listeners can use prosodic cues quickly and effectively to predict upcoming referents (e.g., Ito & Speer, 2008; Weber et al., 2006a). Most of these studies investigate the role of contrastive pitch accents for prediction during discourse processing. Contrastive pitch accents in both English and German consist of a low target, followed by a steep rise in pitch to a high target on the stressed syllable of the pitch-accented word, typically toward the end of the stressed vowel (L + H* using ToBI labeling, cf. Silverman et al., 1992). Thus, they are characterized by a large and salient pitch excursion (as well as lengthening).

Native English listeners can use L + H* accents to predict upcoming referents in a discourse context like *Hang the blue angel. [...] Now, hang the GREEN...* (Ito & Speer, 2008). Specifically, listeners started looking at the angel when hearing *GREEN*, and crucially, before hearing the following noun, thus predicting that the noun *angel* would be repeated. This led to an anticipatory effect if *angel* did indeed follow *GREEN*, but caused a prosodic garden-path effect if a different noun, such as *drum*, followed *GREEN*. Such predictive processing did not occur if the instruction contained no contrastive pitch accent, as in *Hang the blue angel. [...] Now, hang the green...* (Ito & Speer, 2008, Experiment 2).

Similarly, native German listeners looked at a picture of red scissors earlier when hearing German instructions to *Click on the purple scissors* followed by *Click on the RED...* compared to *Click on the red...* (Weber et al., 2006a). Again, facilitation occurred if the previous noun was repeated, and a prosodic garden-path effect occurred if the previous noun was not repeated. Furthermore, there was a smaller garden-path effect for sequences like *Click on the purple scissors. Click on the red vase*. This suggests that participants expected a repeated noun in successive instructions regardless of prosody, but

that a contrastive accent on the adjective of the second instruction strengthened this expectation.

Overall, native English and German listeners can use contrastive pitch accents to predict upcoming referents. Moreover, prosodic information is used rapidly for prediction during processing as the relevant prosodic cue was produced on the word immediately preceding the predicted noun, and prediction thus occurred as soon as listeners heard the cue.

PREDICTIVE PROCESSING IN THE L2

Predictive processing in the L2 differs markedly from the L1 in that L2 learners engage in predictive processing in fewer processing situations than native speakers (e.g., Dussias et al., 2013; Hopp, 2013), and even when their knowledge of the words and syntactic structures involved in the processing is comparable to that of native speakers (e.g., Grüter et al., 2012, 2017; Lew-Williams & Fernald, 2010). Various factors influence whether or not L2 learners engage in predictive processing. For example, more proficient L2 learners can show nativelike predictive processing (Dussias et al., 2013; Hopp, 2013), and there is evidence that L2 learners can engage in more nativelike processing when their L1 is similar to their L2 (Dussias et al., 2013; Foucart & Frenck-Mestre, 2011; Sabourin & Stowe, 2008).

The few studies that have investigated whether and how L2 learners use contrastive pitch accents for predictive processing suggest that L2 learners can use these cues for prediction if they can use knowledge and processing routines from their L1 in their L2. For example, native Spanish L2 learners of English only used contrastive pitch accents predictively in English if equivalent prosodic structures occurred in Spanish (Klassen, 2015). Spanish assigns prominence to the rightmost element in a prosodic phrase and this prominence can shift leftward only in the case of a correction, but not a contrast (Klassen, 2015). Consequently, to indicate a contrast, a prosodic structure like *Move pumpkin number THREE* is possible in Spanish, but a prosodic structure like *Move PUMPKIN number three* is not, whereas both are possible in English. In line with this, native Spanish and native English speakers showed anticipatory eye movements to a picture of pumpkin number two for instructions like *Move pumpkin number THREE to pumpkin number TWO*, which both English and Spanish allow. In contrast, only native English, but not native Spanish, speakers showed anticipatory eye movements to a picture of rocket number three for instructions like *Move PUMPKIN number three to ROCKET number three*, which Spanish does not allow. Similarly, native French speakers, but not native English L2 learners of French, used contrastive prosody to predict upcoming referents in French (Namjoshi, 2015), and native English speakers, but not native Japanese and Chinese L2 learners of English, used contrastive prosody predictively in English (Perdomo & Kaan, 2019; Takeda, 2018). Here, learners could not easily use knowledge and processing routines from their L1 in their L2. Specifically, the intonational systems of French, Japanese, and Chinese differ substantially from that of English. In French, the pitch accent that most typically conveys a contrast is not a L + H* accent, but a high tone (H) on the initial syllable of the contrasted word, which also carries a H* accent on its final syllable (Di Cristo, 1998). Japanese and Chinese prosodically mark contrastive information not through pitch accents, but through local pitch range expansion (Greif, 2010; Venditti et al., 2008).

Overall, the findings so far suggest that L2 learners may only be able to use contrastive pitch accents as a cue for prediction if their native language is sufficiently similar to their L2 in terms of contrastive accentuation.

THE ROLE OF EXPOSURE IN PREDICTIVE PROCESSING

Exposure can influence bilinguals' processing routines, and this effect can be independent of proficiency (Dussias & Sagarra, 2007). Exposure may be especially important for cues like contrastive pitch accents because they are optional, unlike some other cues, such as grammatical gender assignment (Dussias et al., 2013). For example, English adjectives that appear in the context of a lexical contrast receive a contrastive pitch accent that marks this contrast only around 50% of the time. This number goes down even further for lexically contrasted nouns, which receive a contrastive pitch accent only about 20% of the time (Ito & Speer, 2006).

Recent exposure does indeed influence how native listeners interpret contrastive pitch accents (Kurumada et al., 2014, 2012). Specifically, previous exposure to a reliable or unreliable speaker, that is, a speaker whose use of contrastive pitch accents did or did not provide reliable information about referents, influences how native English listeners interpret statements such as *It looks like a zebra* compared to *It LOOKS like a zebra* (Kurumada et al., 2014). Participants exposed to the reliable speaker looked at the picture of an okapi (which has striped legs and looks quite similar to a zebra) reliably more often than at the picture of a zebra when hearing *LOOKS*, but not when hearing *looks*, suggesting that they used the L + H* accent as a cue that the speaker was *not* referring to a zebra. No such effect was found for participants exposed to the unreliable speaker, suggesting that native listeners consider the prior reliability of the prosodic cue when making predictions during language processing.

EXPOSURE AND PREDICTION IN MODELS OF L2 LANGUAGE PROCESSING

Exposure plays a major role in several models of language processing that allow for predictive processing and have been applied to L2 processing. Such models include constraint-based models (Dussias & Cramer Scultz, 2008; MacDonald et al., 1994), tuning models (Cuetos et al., 1996; Dussias & Sagarra, 2007), and implicit learning models (e.g., Chang et al., 2012). Constraint-based models assume that listeners use all the available information immediately during processing and that several alternatives may be activated in parallel. The processor selects an alternative by weighing different constraints, whose strength is determined probabilistically through effects of frequency, plausibility, and so forth (Altmann, 1998). Many of the proposed constraints, such as global syntactic biases, biases of individual words, and so forth, are based on input frequency and are thus directly related to exposure. Similarly, tuning models and implicit learning models assume that processing is experience based, and that listeners would keep track of, for example, whether or not lexical contrasts are prosodically marked with a contrastive pitch accent and would adjust their predictions accordingly (Cuetos et al., 1996). Thus, frequency information derived from exposure plays a major role in these models.

Importantly, all these models also explicitly incorporate predictive processing. Predictions as to what comes next are modeled in terms of weightings or adjustments that are

based on frequency and other information derived from the input. What exactly is predicted depends on the weightings for the individual options, which in turn is influenced by how often each option is encountered in the input and/or by how often each option encountered in the input matches the prediction. Because exposure plays a major role in these models, similar exposure patterns should lead to similar processing.

Some models of L2 language processing do not directly incorporate exposure-based effects and predictive processing into their mechanism, but are compatible with such effects. For example, models that focus on resource deficits, such as computational difficulties in L2 processing (Hopp, 2009; McDonald, 2006), are compatible with both predictive processing and exposure-based effects. These models assume that differences in L1 and L2 processing are mainly due to resource deficits and would predict that L2 learners can generally engage in predictive processing, but may not be able to do so with increasing task complexity, slower lexical access, or less automatic processing routines.

THE CURRENT STUDY

The current study focuses on the role of recent exposure for using contrastive pitch accents to predict upcoming referents in bilinguals' L1 and L2. The specific focus of the current study is on comparing predictive processing across the bilinguals' two languages. That is, the focus is on how participants respond in their L1 versus their L2 when faced with the same changing processing situation.

Experiment 1 compares predictive processing in the L1 and the L2 in intermediate to advanced German–English bilinguals. Based on the previous literature, German–English bilinguals should be able to use their L1 knowledge in their L2 and engage in predictive processing both in the L1 and the L2. However, if participants are generally slower in processing their L2, which would be most compatible with a resource-deficit account, participants may only engage in predictive processing in their L1, but not their L2. During Experiment 1, the speaker that participants encounter uses prosodic cues inconsistently, such that predictive processing should decrease over the course of the experiment (Kurumada et al., 2012, 2014).

Experiment 2 is preceded by an exposure phase in which participants experience the same speaker consistently using contrastive pitch accents as a cue to upcoming referents. This consistent exposure should facilitate predictive processing in the following experimental trials in both the L1 and the L2. During the experimental trials, participants again encounter the speaker using prosodic cues inconsistently, such that predictive processing should again decrease over the course of the experiment, though possibly more slowly than in Experiment 1.

EXPERIMENT 1

METHODS

Participants

Seventeen native-German intermediate-to-advanced (B2 or above using CEFR levels; Council of Europe, 2001) learners of English (4 male, 13 female; mean age 24.5, $SD = 5.2$)

participated in the study. An additional participant was excluded due to more than 20% of track loss. Including this participant in the analysis did not change any of the results. Participants self-rated their English proficiency on a scale with 1 being beginner, 2 being good at English, 3 being very good at English, 4 being fluent, and 5 being native. Participants' average ratings for their reading and writing abilities were 3.3 ($SD=0.8$) and 2.7 ($SD=0.7$), respectively. Comprehension and speaking abilities were rated as an average of 3.0 ($SD=0.7$) and 2.6 ($SD=0.9$), respectively. Participants had been learning English for an average of 10.9 years ($SD=2.9$) at the time of the study.

Materials

The materials for this study comprised line drawings of different objects and recorded instructions to click on these objects. Twenty-four line drawings, which were either under a creative commons license or freely available online, were grouped into sets of four (see Appendix A). In any given trial, six objects from one set were displayed on the computer screen (see Figure 1). The German names for the objects in each set had the same grammatical gender, so that listeners could not identify an object based on hearing the gender-marked definite article that preceded each mention of the object in German (Lew-Williams & Fernald, 2010). Each line drawing was colored in four different colors (blue, green, red, and yellow) using GIMP (The GIMP team, 2014), for a total of 16 objects (four objects in four different colors) in each set.

Instructions to click on the objects were recorded for all objects in all four colors in both German and English by a balanced German–English bilingual with phonetic training. Instructions were of the form *Click on the [color] [object name]*, for example, *Click on the green banana* or *Klick die grüne Banane an* (literally: *Click the green banana on*). All instructions were recorded with three prosodic patterns on the adjective and noun: A L+H* accent on the adjective and no pitch accent on the noun (LHA prosody), no pitch accent on the adjective and a L+H* accent on the noun (LHN prosody), or no L+H*

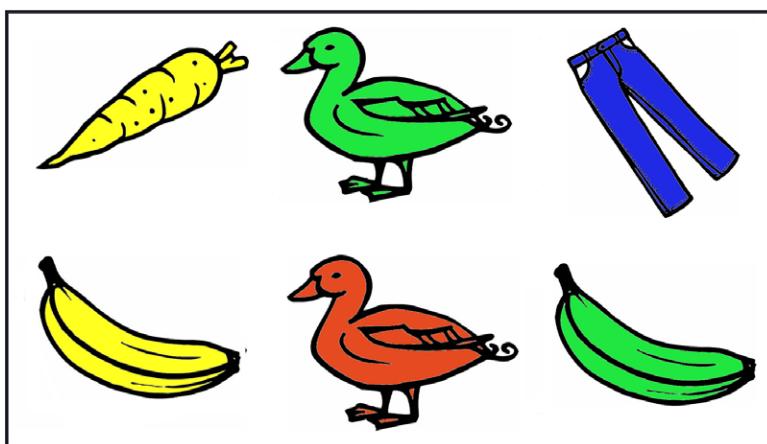


FIGURE 1. Sample experimental display. Objects pictured were adapted from materials by Saskia, Gast, and Janina Valko and are available at madoo.net under a (cc) Creative Commons by-sa license.

accent on the adjective or noun. The latter most typically resulted in a H* accent on the adjective and a !H* on the noun (HH prosody).

Tables 1 and 2 show the means for duration, f0 minimum, f0 maximum, and f0 range (extracted using Praat; Boersma & Weenink, 2017) for the productions in the three prosodic conditions for German and English, respectively. Because this article concerns predictive processing in response to the prosody of the adjective, the table shows these summary measures for the adjective only. F0 minima and maxima were manually checked for pitch halving, doubling, and segmental effects, and were hand corrected if needed. F0 range was calculated by subtracting f0 minimum from f0 maximum. For each of the four measures, Tables 1 and 2 also show the results of a one-way ANOVA analysis and a Tukey's multiple comparisons of means post-hoc test comparing the three prosodic conditions. For the LHA condition, where the adjective carries a L+H* accent, the tables

TABLE 1. Mean duration, f0 minimum, and f0 maximum in the three prosodic conditions (LHA, LHN, and HH) for the adjective in German. Peak location as a percentage of adjective duration is also given for the LHA condition

Prosodic condition	Duration	f0 minimum	f0 maximum	f0 range	Peak location
LHA	484.48 ms (SD = 72.67)	192.13 Hz (SD = 8.43)	303.68 Hz (SD = 21.58)	111.54 Hz (SD = 25.11)	61.39%
LHN	325.59 ms (SD = 44.41)	188.18 Hz (SD = 6.59)	229.32 Hz (SD = 11.91)	45.11 Hz (SD = 12.43)	—
HH	336.40 ms (SD = 49.08)	221.79 Hz (SD = 8.20)	266.91 Hz (SD = 21.00)	41.15 Hz (SD = 19.77)	—
ANOVA	$F = 203.0$ $p < .001$	$F = 534.6$ $p < .001$	$F = 379.6$ $p < .001$	$F = 383.2$ $p < .001$	—
Tukey	$p_{LHA-LHN} < .001$ $p_{LHA-HH} < .001$ $p_{LHN-HH} < .001$	$p_{LHA-LHN} < .01$ $p_{LHA-HH} < .001$ $p_{LHN-HH} < .001$	$p_{LHA-LHN} < .001$ $p_{LHA-HH} < .001$ $p_{LHN-HH} < .001$	$p_{LHA-LHN} < .001$ $p_{LHA-HH} < .001$ $p_{LHN-HH} = .36$	—

TABLE 2. Mean duration, f0 minimum, and f0 maximum in the three prosodic conditions (LHA, LHN, and HH) for the adjective in English. Peak location as a percentage of adjective duration is also given for the LHA condition

Prosodic condition	Duration	f0 minimum	f0 maximum	f0 range	Peak location
LHA	337.88 ms (SD = 58.20)	201.39 Hz (SD = 13.15)	310.63 Hz (SD = 18.49)	109.24 Hz (SD = 12.53)	71.30%
LHN	262.75 ms (SD = 52.55)	174.95 Hz (SD = 6.40)	203.59 Hz (SD = 14.26)	28.63 Hz (SD = 13.87)	—
HH	273.43 ms (SD = 43.80)	213.39 Hz (SD = 7.93)	275.88 Hz (SD = 23.27)	62.49 Hz (SD = 26.64)	—
ANOVA	$F = 59.0$ $p < .001$	$F = 402.3$ $p < .001$	$F = 790.2$ $p < .001$	$F = 445.3$ $p < .001$	—
Tukey	$p_{LHA-LHN} < .001$ $p_{LHA-HH} < .001$ $p_{LHN-HH} = .33$	$p_{LHA-LHN} < .001$ $p_{LHA-HH} < .001$ $p_{LHN-HH} < .001$	$p_{LHA-LHN} < .001$ $p_{LHA-HH} < .001$ $p_{LHN-HH} < .001$	$p_{LHA-LHN} < .001$ $p_{LHA-HH} < .001$ $p_{LHN-HH} < .001$	—

also show the percentage of the adjective that has elapsed when the peak of the contrastive accent is encountered. All the summary measures show significant differences across the three prosodic conditions in both languages, such that the adjective in the LHA condition is, as expected, significantly longer in duration, has a significantly higher f0 maximum, and has a significantly higher f0 range than both the LHN and HH conditions.

An additional two-sample t-test shows that the German target adjectives are significantly longer in duration than the English target adjectives ($t = -15.82$, $df = 344.26$, $p < 0.001$), giving participants more time to process the German compared to the English target adjectives. The possible implications of this will be addressed in the discussion section for Experiment 1.

For each trial, six pictures from one picture set were combined with two recorded instructions to click on two of the displayed objects. For example, the display in Figure 1 was combined with an instruction to click on the red duck and a following instruction to click on the green banana. The two successive instructions either had a repeated noun such that they differed only in the color adjective (Color Contrast Condition), a repeated adjective such that they differed only in object type (Object Contrast Condition), or neither a repeated adjective nor a repeated noun (No Contrast Condition). The first instruction was always produced with rather neutral HH prosody. The second instruction was produced either with LHA prosody (LHA Condition), LHN prosody (LHN Condition), or again with HH prosody (HH Condition). The contrast and prosody conditions were combined to yield eight experimental and filler conditions, listed in Table 3.

The displays that participants saw in the experimental conditions always showed the object mentioned in the first instruction (e.g., a red duck) and two objects in a different color, one of which would be mentioned in the second instruction. Of these, one object had the same object type (e.g., a green duck; Color Contrast condition) and the other a different object type (e.g., a green banana; No Contrast condition) than the first-mentioned object. Moreover, each display contained three filler pictures (e.g., a yellow banana, a yellow carrot, and blue pants, cf. Figure 1). The displays for the filler trials always showed

TABLE 3. Experimental prosodic and contrast conditions

Trial type	Condition	Example
Experimental	Color contrast HH	Click on the red banana. Click on the green banana.
	Color contrast LHA	Click on the red banana. Click on the GREEN banana.
	No contrast HH	Click on the red duck. Click on the green banana.
	No contrast LHA	Click on the red duck. Click on the GREEN banana.
Filler	Object contrast HH	Click on the green duck. Click on the green banana.
	Object contrast HLN	Click on the green duck. Click on the green BANANA.
	No contrast HH	Click on the red duck. Click on the green banana.
	No contrast LNH	Click on the red duck. Click on the green BANANA.

the two objects mentioned in the two instructions (e.g., a red duck and a green banana), each of the two object types in a different color (e.g., a green duck and a yellow banana), and two filler pictures (e.g., a yellow carrot and blue pants).

The order of trials for the German and English versions of the experiment was identical. Picture sets and contrast conditions were distributed across the experiment in a Latin square design. The location of objects from the same picture set differed across trials, so that across the experiment no display was identical. Prosody conditions were distributed so that every other trial had HH prosody, with LHA and LHN prosody distributed pseudorandomly across the experiment.

PROCEDURE

Participants came to the lab on two different days, approximately 1 week apart. The procedure on both days was the same. Half the participants participated in the German version of the experiment in the first session and the English version in the second session, and vice versa for the other half of the participants. After giving informed consent, participants were first familiarized with all the objects and object names through black and white line drawings of the objects on printed cards. The experimenter asked participants to name each object and corrected the object name when needed.

Participants then engaged in two production tasks (first and fourth tasks) and two eye-tracking tasks (second and third tasks), with a short break after each task. The current study focuses on the eye-tracking results, so detailed information about the production tasks will not be reported here. Briefly, participants saw the same kinds of displays and objects as in the eye-tracking tasks (see Figure 1) and produced verbal instructions to click on two successive objects marked as 1 and 2 in the display using the sentence frame *Click on the [color] [object name]*. Thus, the first production task further familiarized participants with the display layout and objects shown in the eye-tracking tasks.

Participants' eye movements were recorded using a Tobii Pro X2-60 remote eye tracker, attached to a Dell 25-inch monitor. Participants were seated at a comfortable distance from the screen and calibrated using a nine-point calibration procedure. Participants were informed that during each trial they would see six different objects on the computer screen and listen to two successive instructions to click on two of the objects. Their task was to follow the instructions and click on the mentioned objects with the computer mouse. Each trial was preceded by 250 ms of blank screen, the first instruction began 200 ms after the onset of the visual display, and the second instruction began 200 ms after participants had clicked on the first object.

For Experiment 1 (second task), participants completed six trials in all the conditions listed in Table 3, for a total of 48 trials. Importantly, this means that the speaker was inconsistent, such that prosody was not informative with respect to contrast condition. For example, *Click on the red duck. Click on the GREEN...* was equally frequently followed by *duck* and *banana*. Participants could thus not develop expectations as to whether or not the noun would be repeated solely based on the prosodic pattern on the adjective.

After a small break, participants completed Experiment 2 (third task), which will be described in more detail in the following text. Afterward, participants completed another production task (fourth task), followed by a language background questionnaire and the opportunity to participate in a gift card drawing.

DATA ANALYSIS

Participants' proportion of looks over time to the target object (numerator), that is, the object mentioned in the second instruction, relative to looks to all six objects shown on the screen (denominator) will be used to measure whether they used contrastive pitch accents to predict upcoming referents in the L1 and the L2. Track loss affected 8.9% of data points, distributed relatively evenly across conditions (difference within 3%). Statistical power is similar to Weber et al. (2006a), which most closely resembles the current study, with 102 trials per condition (17 participants \times 6 trials per condition) compared to Weber et al.'s 96 (24 participants \times 4 trials per condition).

If participants use L + H* accents predictively, they should expect a repeated noun upon encountering a L + H*-accented adjective, but not upon encountering an adjective with no L + H* accent. Thus, for the LHA conditions, the curves that show proportion of looks over time should rise earlier in the Color Contrast condition compared to the No Contrast condition. But for the HH conditions, the proportion of looks over time curves should rise at a similar time in the Color Contrast and No Contrast conditions.

Of particular interest for the current study is at which points in time participants look significantly more at the target object in the Color Contrast condition compared to the No Contrast condition. This will be investigated using Smoothing Spline ANOVA analyses (SSANOVA; cf. Gu, 2013), a statistical analysis used to compare curves. SSANOVAs allow for a holistic comparison of curves and can tell us when over time a particular condition yields significantly more looks to the target object over other conditions (Davidson, 2006; Gu, 2013). To do this, SSANOVAs fit smoothing splines to the curves of the experimental conditions being compared. The original eye-tracking data is typically rather noisy and produces jagged curves. Smoothing splines determine which smoothed curves best fit the data by balancing goodness-of-fit and smoothness of the original curve. Bayesian confidence intervals can then tell us which sections of the curves diverge statistically significantly. Specifically, we find statistically significant differences between two curves where their confidence intervals do not overlap (Chanethom, 2011; Koops, 2010). This, in turn, tells us when over time participants have reliably more looks to the target object in one condition compared to another. Thus, significance can be determined visually from the plotted curves and confidence intervals. Data and analysis scripts are available on the Open Science Framework at <https://osf.io/vuwq9>

RESULTS

L1 PROCESSING

Figure 2a shows participants' proportion of looks over time to the target object in the target conditions in the L1. All results figures are aligned such that 0 ms on the x-axis, visually represented by a vertical solid line, represents the end of the adjective, which coincides with the beginning of the noun. Because it takes about 150 ms–200 ms to plan and execute an eye movement (Fischer, 1998), an additional, vertical dashed line at 150 ms shows the earliest point in time at which we would expect eye movements to the target object in response to hearing the beginning of the noun. Thus, all eye movements to the target object that occurred before 150 ms were clearly in response to hearing the adjective of the instruction, and before disambiguating segmental information from the noun had arrived.

[Figure 2a](#) shows that looks to the target object started rising earliest for a L + H* accent on the adjective followed by a repeated noun and last for a L + H* accent on the adjective followed by a different noun. This pattern mirrors the results from previous studies with native German speakers.

A SSANOVA investigated when in time the four curves diverge significantly. The SSANOVA was conducted in R using the gss package (Gu, 2014) and the gssanova () function, which fits SSANOVA models for non-Gaussian responses. Because responses from the current eye-tracking experiment are binomial (0 = not looking at the target; 1 = looking at the target), a binomial error distribution was selected (`family = "binomial"`). The statistical model's response variable was looks to the target (0 vs. 1), and the fixed factors were time (from -500 to 1,000), contrast condition (Color Contrast vs. No Contrast), prosody condition (LHA vs. HH), and all their interactions. The model also included random intercepts for participant and item. All SSANOVA graphs in this article show curves derived from analyses containing all fixed factors and interactions. However, the HH and LHA conditions are plotted separately here and in all the following SSANOVA graphs because graphs showing curves and confidence intervals for all four conditions would be rather cluttered and difficult to read. Thus, [Figure 2b](#) shows the smoothed curves and the 95% Bayesian confidence intervals for the HH conditions, whereas [Figure 2c](#) shows them for the LHA conditions.

[Figure 2b](#) shows that the modeled looks to the target object start rising earlier in the Color Contrast HH Condition compared to the No Contrast HH Condition, but at no point in time do the two curves diverge sufficiently for this difference to reach significance. Thus, there is no evidence that participants predict a repeated noun if the adjective does not receive a L + H* accent. [Figure 2c](#) shows that the modeled looks to the target object start rising earlier in the Color Contrast LHA Condition compared to the No Contrast LHA Condition, and that the curves diverge significantly from around 0 ms to about 650 ms. Thus, participants start looking at the target object in the LHA conditions earlier when the noun is repeated compared to when it is not. Importantly, this difference is significant in a window that starts before 150 ms, that is, during the processing of the color adjective. This suggests that participants are predicting a repeated noun rather than a different noun if the adjective received a L + H* accent.

To explore whether exposure to an inconsistent speaker over the course of the experiment influences predictive processing in the L1, [Figure 3](#) shows the SSANOVA results for the LHA conditions separately for the first and second halves of the experiment. [Figure 3a](#) shows that during the first half of the experiment, participants engage in predictive processing, with the Color Contrast and No Contrast curves diverging from about -50 ms to 750 ms. [Figure 3b](#) shows results from the second half of the experiment and suggests that, as participants are exposed to the inconsistent speaker over the course of the experiment, their processing is no longer predictive. That is, the Color Contrast and No Contrast curves diverge reliably only from 200 ms to 450 ms, that is, when segmental information from the noun has started arriving.

L2 PROCESSING

[Figure 4a](#) shows participants' proportion of looks over time to the target object for the experimental conditions in the L2. The figure shows the same overall pattern as [Figure 2a](#)

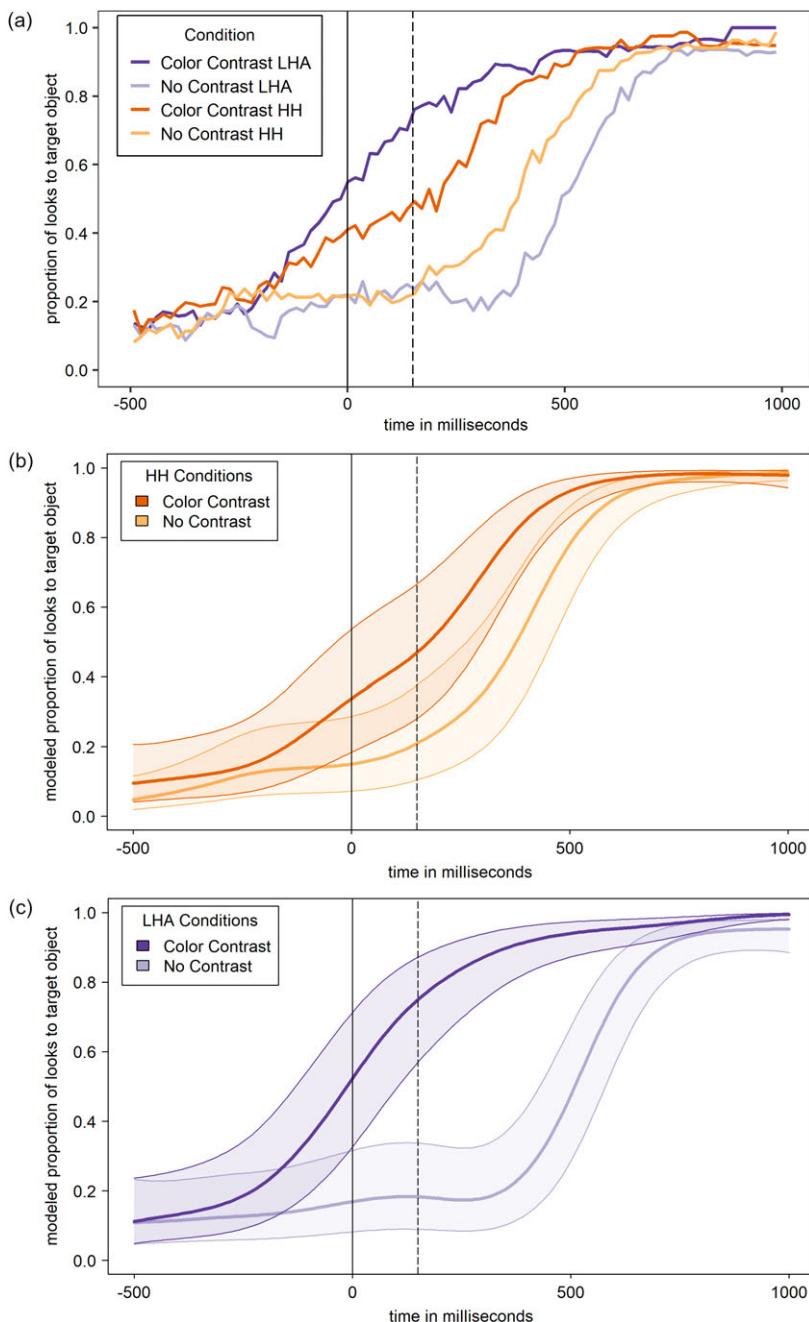


FIGURE 2. (a) Proportion of looks over time to the target object across target conditions in German (L1). SSANOVA results comparing looks to the target object for (b) HH and (c) LHA conditions.

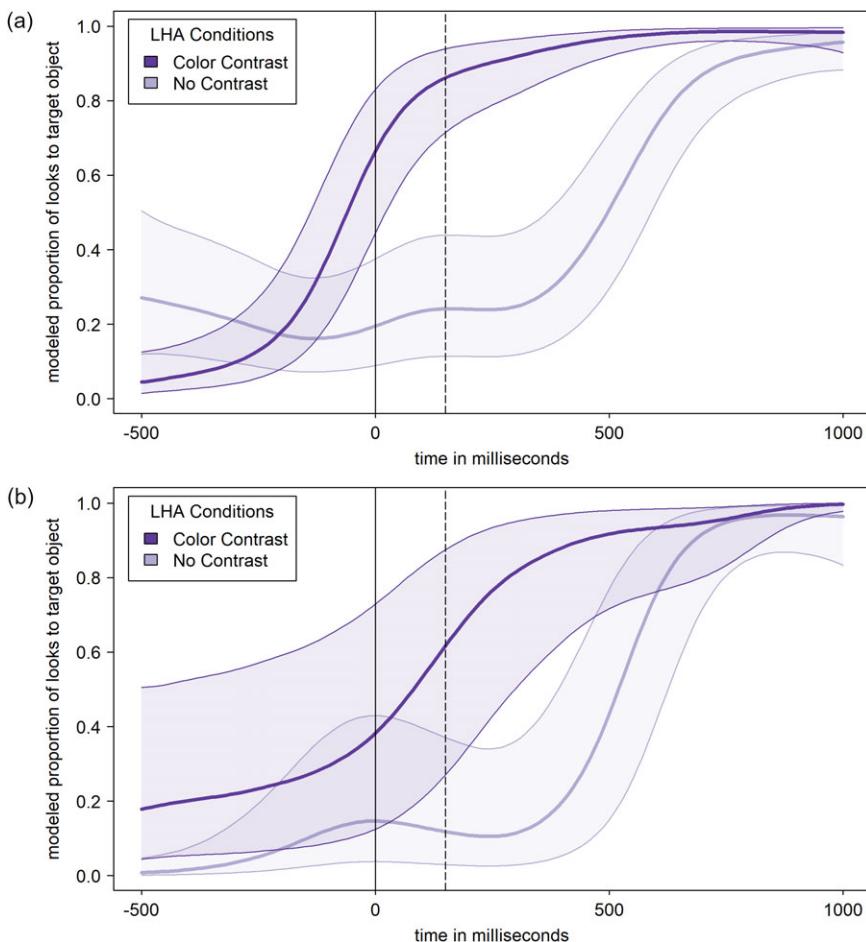


FIGURE 3. SSANOVA results for German (L1) comparing looks to the target object for LHA conditions in the (a) first and (b) second half of the experiment.

for the L1 data. However, the curves cluster much closer together than in the L1 data. A SSANOVA identical to the one described in the preceding text investigated when in time these four curves diverge significantly. Figures 4b (HH conditions) and 4c (LHA conditions) show the smoothed curves and the 95% Bayesian confidence intervals.

Figure 4b shows that the modeled looks to the target object start rising earlier in the Color Contrast HH Condition compared to the No Contrast HH Condition. Similar to the L1 data mentioned previously, at no point in time do the two curves diverge statistically significantly. Figure 4c shows that the modeled looks to the target object again start rising earlier in the Color Contrast LHA Condition compared to the No Contrast LHA Condition, and the curves diverge significantly from around 275 ms to 600 ms. Thus, participants started looking at the target object in the LHA conditions earlier when the noun is repeated compared to when it is not. However, the window starts after 150 ms, that is, after the color adjective had been fully processed and participants have started hearing the

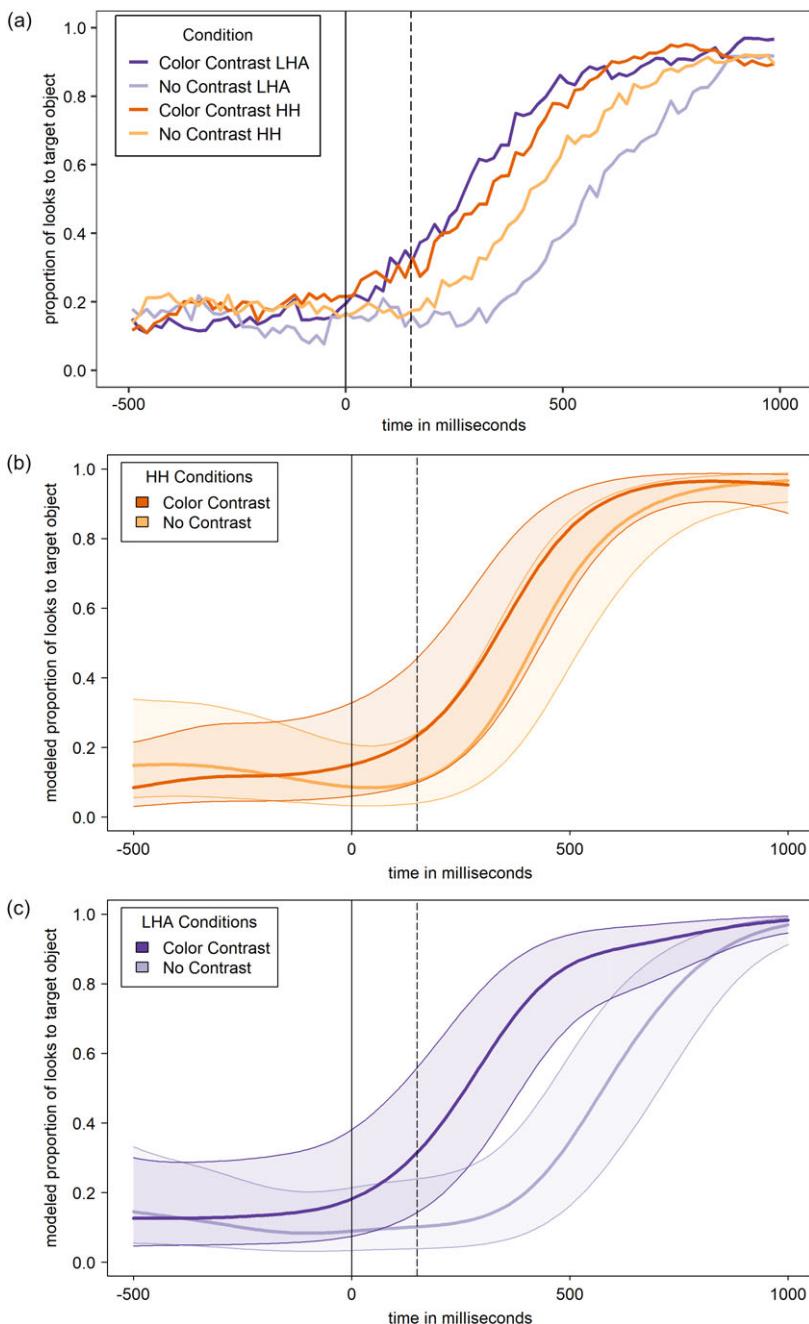


FIGURE 4. (a) Proportion of looks over time to the target object across target conditions in English (L2). SSANOVA results comparing looks to the target object for (b) HH and (c) LHA conditions.

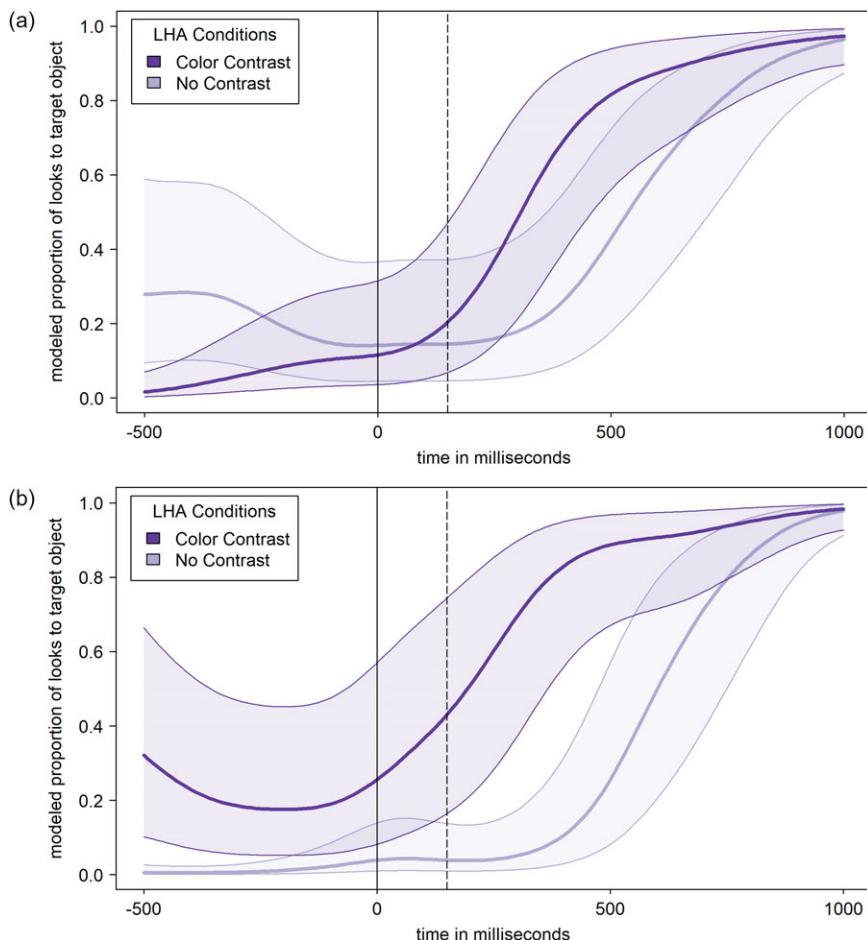


FIGURE 5. SSANOVA results for English (L2) comparing looks to the target object for LHA conditions in the (a) first and (b) second half of the experiment.

beginning of the target noun. This suggests that participants show facilitative processing for a repeated noun compared to a different noun if the adjective received a L + H* accent, but there is no evidence that this facilitative processing is predictive.

To explore whether exposure to an inconsistent speaker over the course of the experiment influences predictive processing in the L2, Figure 5 shows the SSANOVA results for the LHA conditions separately for the first and second halves of the experiment. Figure 5a shows that Color Contrast and No Contrast curves do not diverge during the first half of the experiment, suggesting participants initially do not engage in predictive processing. Figure 5b shows results from the second half of the experiment and suggests that, as participants are exposed to the inconsistent speaker over the course of the experiment, their processing starts showing evidence for being predictive. Specifically, the Color Contrast and No Contrast curves diverge reliably from about 125 ms to 525 ms, that is, the curves start diverging just slightly before segmental information from the noun has started arriving.

DISCUSSION

Experiment 1 investigated whether German–English bilinguals used contrastive pitch accents predictively in both their L1 and L2, and whether their processing changed over the course of the experiment. Participants showed an overall advantage in processing repeated nouns following a L + H* accent in both their L1 and L2. This suggests that participants could clearly use the prosodic cue in both languages. However, while this advantage was clearly a result of predictive processing in the L1, with the advantage occurring before 150 ms, this was not the case for the L2. Here, the advantage occurred after segmental information from the beginning of the noun had come in. A close look at the stimuli suggests that this may be due to the time that participants had to process L + H*-accented adjectives in their L1 German versus their L2 English. The bisyllabic German L + H*-accented adjectives were significantly longer compared to the mostly monosyllabic English L + H*-accented adjectives. Furthermore, the peak of the L + H* occurred on average 85 ms earlier in the German L + H*-accented adjectives compared to the English L + H*-accented adjectives (187 ms compared to 102 ms before the beginning of the disambiguating noun; two-sample t-test: $t = -14.10$, $df = 166.53$, $p < 0.001$). Thus, participants had about 85 ms more time to predict upcoming referents based on the prosodic cues in their L1 German compared to their L2 English. However, even when adjusting for these timing differences in the stimuli, participants in Experiment 1 would still exhibit nonpredictive processing in their L2. Specifically, if the L2 curves were adjusted by 85 ms, they would diverge at around 190 ms, that is, once segmental information from the noun has started arriving. Furthermore, monolingual native English speakers show evidence of clear predictive processing in their L1 English with comparable stimuli and in a somewhat more complex discourse situation (Ito & Speer, 2008). What remains is evidence for predictive processing in the L1 and no such evidence for the L2.

This pattern of results supports resource-deficit accounts over exposure-based accounts for several reasons. Resource-deficit accounts assume no fundamental differences between L1 and L2 processing, but attribute observed differences to resource limitations. Participants' L1 processing seems to differ from their L2 processing not fundamentally, but in terms of processing speed. Specifically, participants use prosodic cues in both their L1 and L2 in a similar manner, but more slowly in the L2, such that we find evidence for predictive processing in their L1, but not in their L2. These differences could be due to slower lexical access in the L2 or less automatic processing routines. Notice that participants' L1 and L2 processing differs even though their two languages use the same prosodic cue to mark a lexical contrast, which would allow participants to use knowledge and processing routines from their L1 in the L2.

The evidence from the first and second halves of the experiment suggest that participants adjusted their predictions in the direction expected by exposure-based accounts in the L1, but not the L2. Exposure-based accounts would predict that participants should engage in less predictive processing over the course of the experiment because the speaker's use of prosodic cues was inconsistent. Moreover, listeners were exposed to trials in which they experienced a prosodic garden-path effect, namely, when hearing a sequence like *Click on the red duck. Click on the GREEN banana*. Such a pattern is not only infelicitous, it also almost never occurs in natural production data (Ito & Speer, 2006). From a constraint-based perspective, these garden-path trials should lead to a large

prediction error, that is, a large difference between what participants predicted and what is then encountered. Such prediction errors should lead to an adjustment of the predictions, such that participants should be less and less likely to use contrastive pitch accents to predict a repeated noun over the course of experimental trials. This happened in the L1, but not the L2, where participants engaged in more rather than less predictive processing over the course of the experiment.

There are several possible explanations for the unexpected L2 findings, and further studies are clearly needed to determine which mechanisms underlie the L2 patterns found here. One possibility is that participants did not show predictive processing in the initial half of the experiment because of resource deficits. Specifically, participants may initially have been substantially slower in their processing, especially in terms of processing routines and/or lexical access. Such slower processing would explain why there is no evidence for predictive L2 processing in the first half of the experiment. As participants get used to the task and the lexical items used, their lexical access and processing may speed up, yielding the reliable predictive processing in the second half of the experiment. Such an explanation would also entail that participants were less sensitive in the L2 than in the L1 to the speaker's prosodic cues being inconsistent. Otherwise, their L2 predictive processing should nevertheless have decreased over the course of the experiment, not increased. It is possible that participants' processing resources in the L2 are taken up by processes such as lexical access and developing processing routines, so that fewer resources are available to track the consistency of the speaker's prosodic cues.

EXPERIMENT 2

Experiment 2 investigates how a brief exposure phase with highly consistent prosodic cues influences predictive processing in bilinguals.

METHODS

Participants, Materials, Procedure, and Data Analysis

Participants, materials, experimental design constraints, procedure, and data analysis were the same as in Experiment 1, except that an exposure phase with an additional 24 trials preceded the 48 experimental trials. During the exposure phase participants experienced the speaker using consistent prosodic cues and heard only felicitously placed L + H* accents, with 12 No Contrast HH trials, 6 Color Contrast LHA trials, and 6 Object Contrast LHN trials. Importantly, this means that instructions like *Click on the red duck. Click on the GREEN...* were always followed by a repeated noun, that is, *duck*, and instructions like *Click on the red duck. Click on the green...* always preceded a different noun, for example, *banana*. This highly consistent cue should allow participants to develop expectations as to whether or not the noun would be repeated based on the prosodic pattern on the adjective. As in Experiment 1, the speaker was inconsistent and prosody was not informative with respect to contrast condition during the experimental trials (cf. conditions in Table 3). Data analysis was the same as in Experiment 1, with track loss affecting 11.7% of data points, again distributed evenly across conditions (difference within 3%).

RESULTS

L1 PROCESSING

[Figure 6a](#) shows participants' proportion of looks over time to the target object in the experimental conditions following the exposure phase in the L1. The overall pattern shown in [Figure 6a](#) mirrors the results from previous studies with native German speakers and the results from Experiment 1. The same SSANOVA analysis as in Experiment 1 investigated when in time the four curves diverge significantly. [Figure 6b](#) shows that while participants start looking at the target object in the HH conditions slightly earlier when the noun is repeated compared to when it is not, this difference does again not reach significance. [Figure 6c](#) shows that participants start looking at the target object earlier in the Color Contrast LHA Condition compared to the No Contrast LHA Condition. This is significant in a large window that, importantly, starts before 150 ms.

Again, to explore whether exposure to an inconsistent speaker over the course of the experiment influences predictive processing, [Figure 7](#) shows the SSANOVA results for the LHA conditions in the L1 separately for the first and second halves of the experiment. [Figure 7](#) show that participants engage in predictive processing in both the first and second halves of the experiment, with the Color Contrast and No Contrast curves diverging from about 25 ms to 1,000 ms in [Figure 7a](#) and from about -5 ms to 1,000 ms in [Figure 7b](#).

L2 PROCESSING

[Figure 8a](#) shows participants' proportion of looks over time to the target object for the experimental conditions in the L2. The figure shows the same overall pattern as [Figure 6a](#) for the L1 data. As in Experiment 1, the curves for the L2 cluster much closer together than those for the L1, but the curve for the Color Contrast LHA Condition seems to be more strongly separated from the other three curves than in Experiment 1. A SSANOVA identical to the one in Experiment 1 investigated when in time the four curves diverge significantly. [Figures 8b](#) and [8c](#) show the smoothed curves and the 95% Bayesian confidence intervals.

[Figure 8b](#) shows that, similar to [Figures 2b](#), [4b](#), and [6b](#), the modeled looks to the target object start rising earlier in the Color Contrast HH Condition compared to the No Contrast HH Condition, but at no point in time do the two curves diverge statistically significantly. [Figure 8c](#) shows that the modeled looks to the target object start rising earlier in the Color Contrast LHA Condition compared to the No Contrast LHA Condition. Importantly, the curves start diverging significantly as early as 50 ms and thus reflect predictive processing.

To explore whether exposure to an inconsistent speaker over the course of the experiment influences predictive processing, [Figure 9](#) shows the SSANOVA results for the LHA conditions in the L2 separately for the first and second halves of the experiment. [Figure 9a](#) shows that the Color Contrast and No Contrast curves do not diverge significantly during the first half of the experiment, suggesting that participants again do not initially engage in predictive processing. [Figure 9b](#) shows results from the second half of the experiment and suggests that, as participants are exposed to the inconsistent speaker over the course of the experiment, their processing starts showing evidence for being predictive. Specifically, the Color Contrast and No Contrast curves diverge reliably from about 125 ms, that is, just before segmental information from the noun has started arriving.

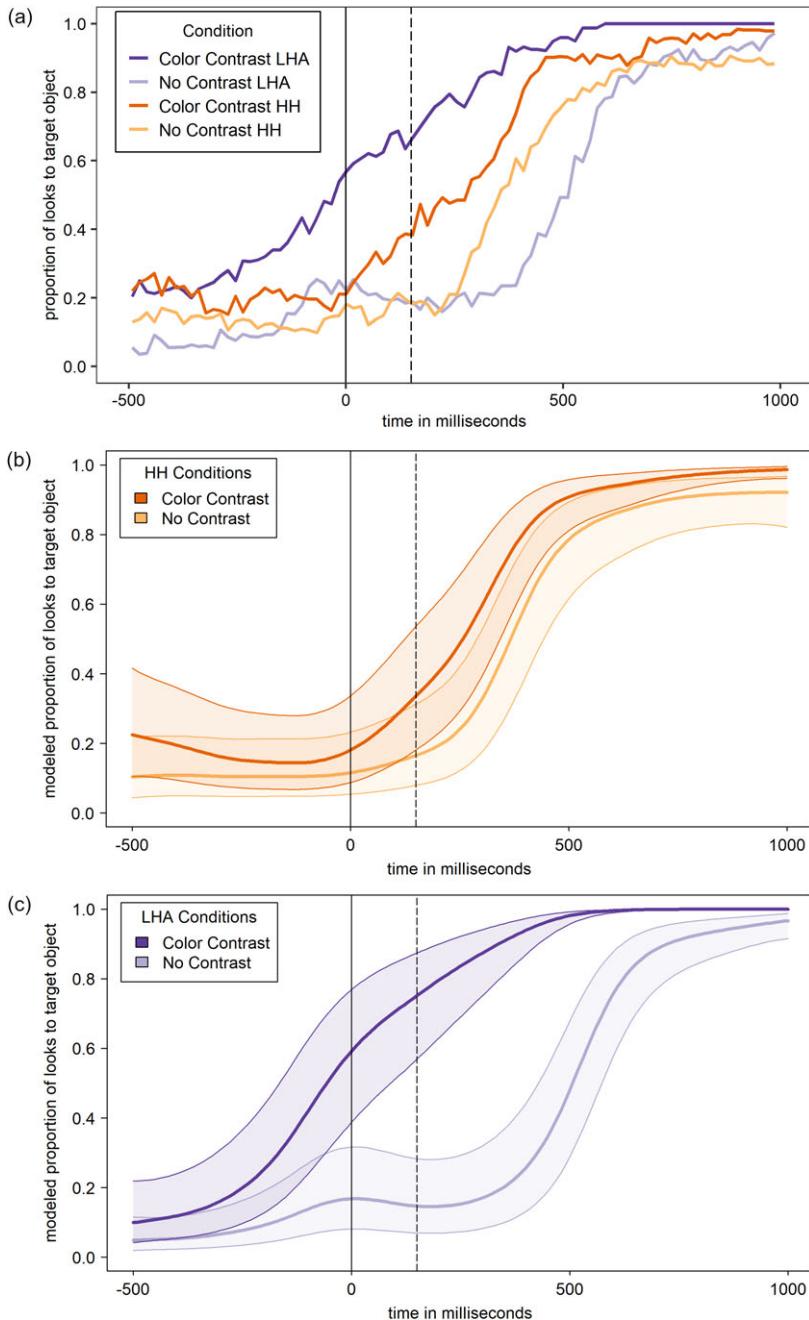


FIGURE 6. (a) Proportion of looks over time to the target object across target conditions in German (L1) after consistent exposure. SSANOVA results comparing looks to the target object for (b) HH and (c) LHA conditions after consistent exposure.

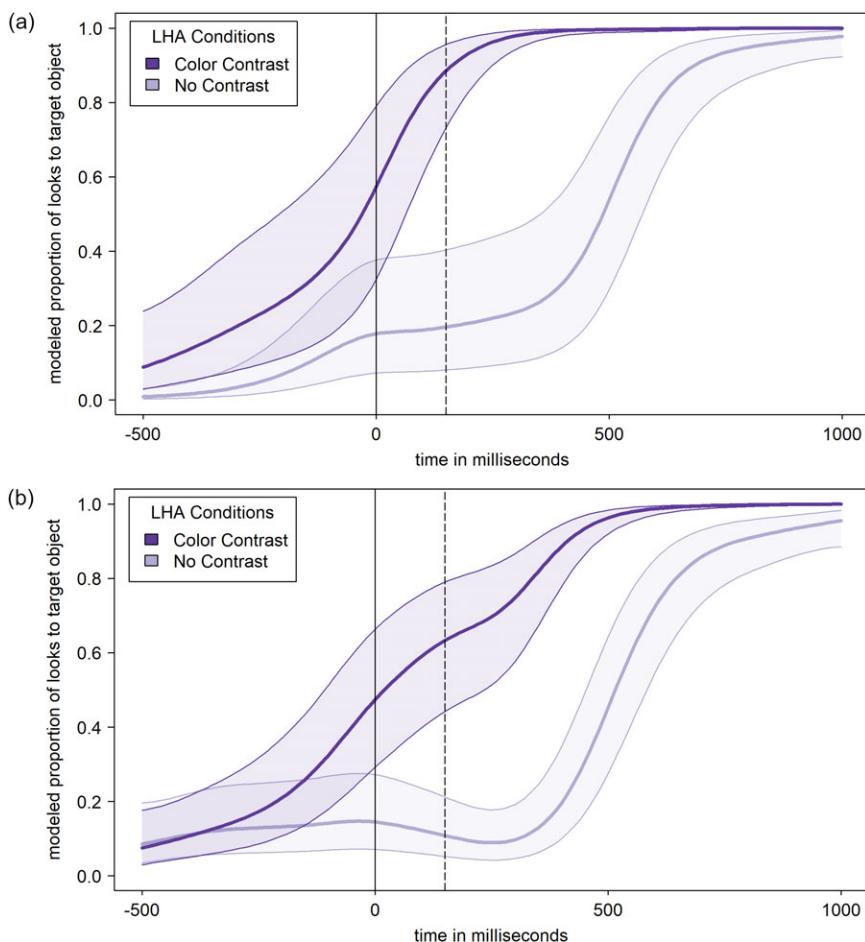


FIGURE 7. SSANOVA results for German (L1) comparing looks to the target object for the LHA conditions after consistent exposure in the (a) first and (b) second half of the experiment.

DISCUSSION

Experiment 2 investigated German–English bilinguals’ predictive processing in both their L1 and L2 following brief recent exposure to a highly reliable speaker. The results suggest that, for the experiment as a whole, participants can indeed use prosodic cues in both their L1 and L2 to engage in predictive processing.

The evidence from the first and second halves of the experiment suggest that exposure-based accounts can again explain the L1 patterns, but not the L2 patterns. Specifically, with previous prosodically consistent exposure, participants did not decrease their predictive processing in the L1 when the speaker was no longer consistent in their use of prosodic cues. Exposure-based accounts are compatible with this result: predictive processing when exposed to an inconsistent speaker should decrease more gradually following exposure to a highly consistent speaker.

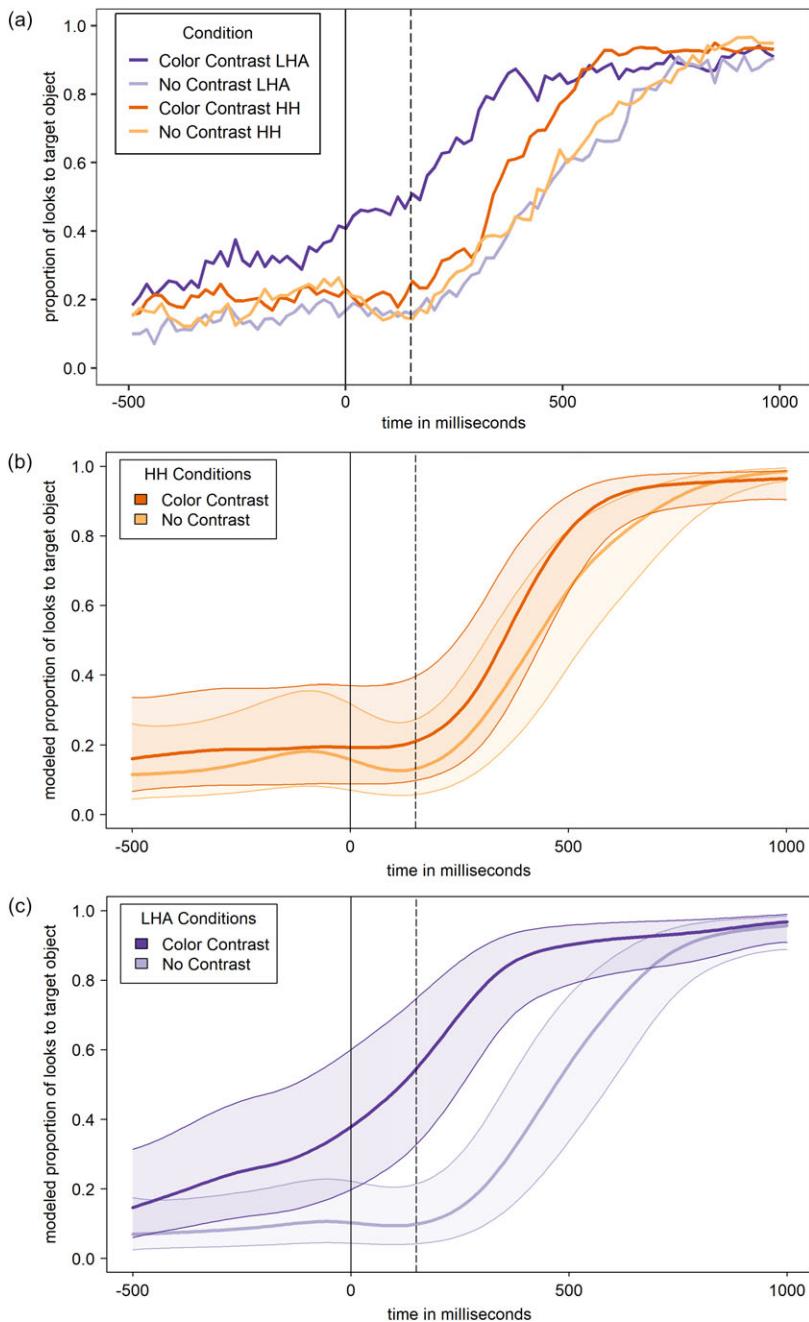


FIGURE 8. (a) Proportion of looks over time to the target object across target conditions in English (L2) after consistent exposure. SSANOVA results comparing looks to the target object for (b) HH and (c) LHA conditions after consistent exposure.

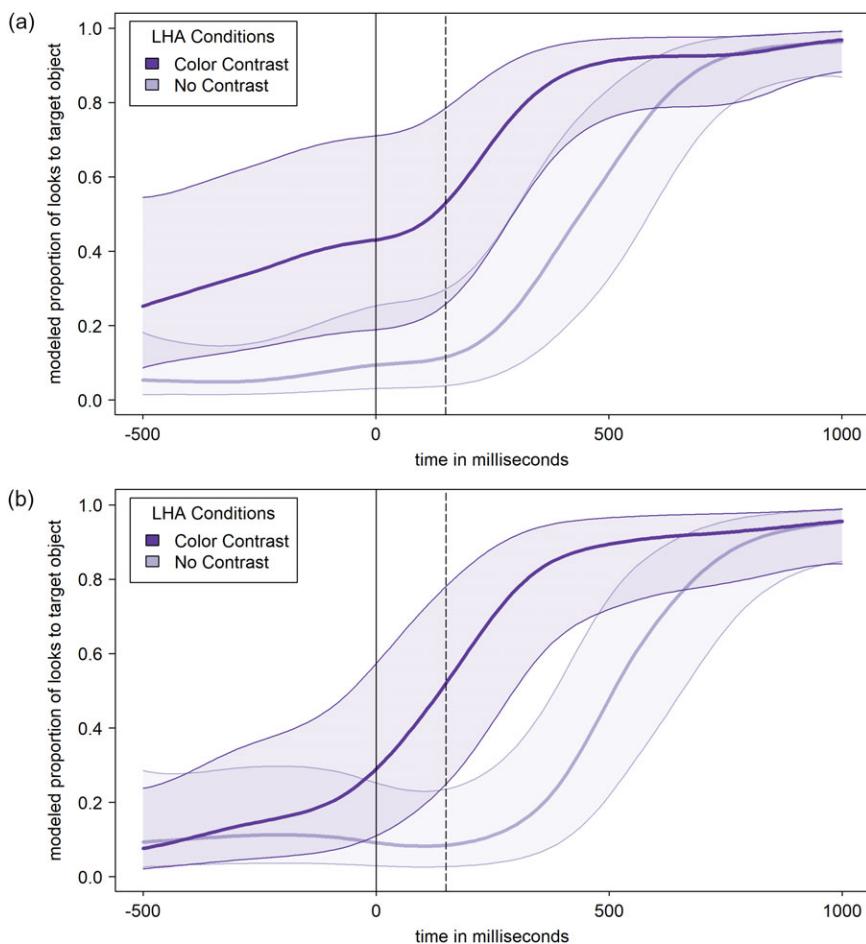


FIGURE 9. SSANOVA results for English (L2) comparing looks to the target object for the LHA conditions after consistent exposure in the (a) first and (b) second half of the experiment.

Again, the L2 findings were unexpected in that participants engaged in more rather than less predictive processing over the course of the experiment, such that further studies are needed to understand the processes involved in L2 processing of prosodic cues. In fact, the L2 results from the first and second halves of Experiment 2 are even more puzzling than those of Experiment 1. Specifically, participants completed Experiment 2 immediately after Experiment 1 and should therefore have been used to the task and the lexical items used. This should have resulted in lexical access and processing that is sufficiently speedy to find evidence for prediction. Despite this, participants initially did not engage in predictive processing. One possibility is that there are individual differences in how participants responded to the speaker changing from being consistent in their use of prosodic cues in the exposure phase to being inconsistent in the first half of the experiment. The large Bayesian confidence intervals in Figure 9a up to about 150 ms, that is, before processing the final noun, support this idea.

GENERAL DISCUSSION AND CONCLUSIONS

PREDICTIVE PROCESSING IN THE L1

The current results show that German–English bilinguals used contrastive pitch accents to predict upcoming referents in their L1 German both before and after the recent exposure session. These results are in line with Ito and Speer (2008), Weber et al. (2006a), and other studies showing that native listeners can use contrastive pitch accents to predict upcoming referents. This process is very efficient and fast: Before exposure participants started using this information in their L1 about 150 ms before the end of the cue-bearing adjective, and thus less than 50 ms after encountering the peak on the L + H* accent. After exposure, participants started using this information in their L1 about 200 ms before the end of the cue-bearing adjective, and thus on average before they encountered the peak of the L + H* accent. Furthermore, the results from the first and second halves of the experiment are compatible with exposure-based accounts of L1 language processing, such that exposure to an inconsistent speaker decreases participants' engagement in predictive processing over time and previous exposure to a highly consistent speaker delays such a decrease.

PREDICTIVE PROCESSING IN THE L2

The L2 results from Experiment 1 are in line with the previous literature on predictive L2 processing, which suggests that prediction occurs in fewer processing situations in the L2 than in the L1 (Kaan, 2014). Similar to Perdomo and Kaan (2019), Experiment 1 found evidence for facilitation, but not for predictive processing in the L2.

Notably, the results from Experiment 1 are not compatible with the idea that bilinguals can engage in predictive processing only if they can use their L1 knowledge in their L2. The native German bilinguals in the current study did not engage in predictive processing in their L2 English even though both German and English use the same pitch accent to convey a contrast (L + H*), and even though the very same participants used the very same cue successfully for prediction in their L1. In this respect, the current results are inconsistent with Klassen's (2015) result for Spanish–English bilinguals. The reason for this discrepancy is most likely timing (Perdomo & Kaan, 2019). In Klassen's (2015) study, the two phrases of an instruction were separated by 700 ms of silence, which served as the analysis window for statistical analysis. Thus, participants had 700 ms to engage in predictive processing and look at the possible target picture after hearing *Move pumpkin number THREE* and before the instruction continued as *to pumpkin number TWO*. This gave L2 learners ample time to process the contrastive pitch accent and generate predictions about upcoming referents before these referents were mentioned. In contrast, in the current study the target noun immediately followed the cue-bearing adjective.

The current results suggest that L2 learners can engage in predictive processing in the L2 if they can use knowledge from their L1 in their L2 and if there is enough time to process the prosodic cue and generate predictions. The study also suggests that prediction takes considerably longer in the L2 than in the L1, such that whether or not L2 learners engage in predictive processing in the L2 partially depends on how much time they have to process the cue and generate predictions before identifying or disambiguating information arrives. What is still an open question is what kind of knowledge from the L1 is

needed to successfully predict, that is, in which ways the L1 and L2 need to be similar, and whether such similarity is critical or merely advantageous for prediction in L2 processing.

EXPOSURE TO CONSISTENT AND INCONSISTENT PROSODIC CUES

The overall results from Experiment 2 show that, after a brief exposure period during which the speaker is highly consistent, participants engaged in predictive processing not just in their L1 but also in their L2. These findings support the idea that exposure to a consistent speaker can increase listeners' reliance on the cues that this particular speaker uses (Kurumada et al., 2012). However, the results from the first and second halves of both Experiments 1 and 2 paint a different picture. While exposure to the consistent or inconsistent speaker over time modulates participants' predictive processing in the L1 in ways that are compatible with exposure-based accounts of language processing, this is not the case for predictive processing in the L2. Specifically, participants showed increased predictive processing in their L2 the more they were exposed to an inconsistent speaker during experimental trials. These results are not easily integrated into exposure-based models or resource-deficit models. Overall, the picture that emerges is one of participants who are quite sensitive to distributional frequencies in the input and who respond to these frequencies in an expected manner in their L1. However, participants seem to respond to the distributional frequencies differently in their L2 than their L1, and what drives these differences is not clear. Specifically, it is not clear whether these differences derive from differences in participants' L1 and L2 processing mechanisms or from the nature and retrieval of the information used by the processing mechanisms.

THEORETICAL IMPLICATIONS

Overall, the current results provide evidence for resource-based over exposure-based accounts of L2 processing. Specifically, resource limitations can more readily explain the absence of predictive processing in the L2 in Experiment 1. Both German and English use the same prosodic cues to mark lexical contrasts and monolingual native speakers of both languages use these cues to engage in predictive processing. An exposure-based account would thus predict that the German–English bilinguals would be exposed to the same prosodic cues regardless of whether they are exposed to German or English, and as a result would not differ in their processing of these prosodic cues in German or English. However, Experiment 1 found a clear delay in processing in the L2 compared to the L1, which can easily be explained in terms of resource limitations, such as slower lexical access and less automated processing, but not in terms of exposure.

Despite this, exposure clearly plays a role in L2 processing. More specifically, it is recent exposure to a highly consistent speaker that allowed for speedier processing in the L2 and thus similar L1 and L2 processing overall in Experiment 2. Furthermore, it is probably long-term exposure to these prosodic cues that allow participants to engage in facilitatory processing in both languages to begin with. Thus, the picture that emerges from the current experiment is of a resource-limitation approach to processing contrastive pitch accents in the L2, where recent exposure to a highly consistent speaker can mitigate resource limitations and allow for more nativelike processing.

In contrast, recent exposure to highly *inconsistent* prosodic cues results in different L1 and L2 responses in terms of processing. It is currently not clear what drives these differing L1 and L2 responses to an inconsistent speaker, and whether the patterns found point to quantitative or qualitative differences in L1 and L2 predictive processing. The idea that L2 processing may be qualitatively different from L1 processing has mainly been discussed for syntactic processing and has been quite influential in this domain. The shallow structure hypothesis claims that native speakers and L2 learners differ fundamentally in how they process certain kinds of complex sentences (Clahsen & Felser, 2006, but see Clahsen & Felser, 2018). Specifically, the shallow structure hypothesis assumes that L2 learners may engage in more shallow processing and may build less detailed syntactic representations for these complex sentences, thus assuming a qualitative difference between L1 and L2 syntactic processing. In the domain of prosody, Lee and Fraundorf (2017) propose a shallow representation account of processing contrastive pitch accents. This account proposes that L2 listeners may consider the salient alternative when processing contrastive pitch accents, that is, consider that the two successively mentioned objects differ only in color and thus consider a repeated noun, but may not fully integrate this alternative into memory because they lack sufficient cognitive resources when processing in their L2. The account further assumes quantitative rather than qualitative differences between L1 and L2 processing, with the quantitative differences being due to resource limitations. Similarly, the Interface Hypothesis (Sorace, 2011) suggests that learners may be less efficient at integrating information across interfaces, in this case prosodic and lexical information, possibly due to processing limitations. Finally, resource limitations may be due to the interpretation of contrastive accents being cognitively demanding. Specifically, listeners need to maintain a set of referents in memory when making predictions about upcoming referents based on contrastive pitch accents (Reichle & Birdsong, 2014). For example, for a sequence of instructions to click on the red duck and then the green duck, participants would have to remember both the previous color and the previous object (i.e., *red* and *duck*) to link a contrastive accent on *GREEN* to the previous noun *duck* and expect a repetition of that particular noun. While bilinguals clearly respond to the inconsistent speaker differently in terms of predictive processing in their L1 and L2, neither the shallow structure hypothesis' nor the shallow representation accounts' approach would predict that L2 speakers would engage in *more* predictive processing over time when faced with inconsistent prosodic cues. That said, it is possible that the patterns observed here are related to shallow processing of some kind, but it is not clear which aspect of L2 processing would be shallow in the current context and whether this kind of shallow processing should be considered to be qualitatively or quantitatively different from L1 predictive processing.

CONCLUSIONS

The current study explored the role of recent exposure in the processing of contrastive pitch accents in German–English bilinguals. The results suggest that even if bilinguals can use information and processing routines from their L1, their L2 processing may still be delayed compared to their L1 processing. Moreover, bilinguals adjust their predictive processing differently in their L1 and L2 when exposed to an inconsistent speaker. Overall, the results support exposure-based accounts for L1 processing, but resource-deficit accounts for L2 processing, with exposure to highly consistent cues possibly mitigating resource deficits.

REFERENCES

- Altmann, G. T. (1998). Ambiguity in sentence processing. *Trends in Cognitive Sciences*, 2, 146–152.
- Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer* (Version 6.0.36) [Computer software]. <http://www.praat.org/>
- Boland, J. E. (2005). Visual arguments. *Cognition*, 95, 237–274.
- Chanethom, V. (2011). Dynamic differences in child bilinguals' production of diphthongs. In *Proceedings of the 4th ISCA Workshop on Experimental Linguistics* (pp. 55–58).
- Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass*, 6, 259–278.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42.
- Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, 40, 693–706.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. E. García-Albea, & N. Sebastián-Gallés (Eds.), *Language processing in Spanish* (pp. 145–187). Lawrence Erlbaum Associates.
- Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America*, 120, 407–415.
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8, 1117–1121.
- Di Cristo, A. (1998). Intonation in French. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 195–218). Cambridge University Press.
- Dussias, P. E., & Cramer Scultz, T. R. (2008). Spanish-English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica*, 128, 501–513.
- Dussias, P. E., & Saggar, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 10, 101–116.
- Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35, 353–387.
- Fischer, B. (1998). Attention in saccades. In R. D. Wright (Ed.), *Visual attention* (pp. 289–305). Oxford University Press.
- Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, 14, 379–399.
- Greif, M. (2010). Contrastive focus in Mandarin Chinese. *Proceedings of Speech Prosody*, paper 836.
- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28, 191–215.
- Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, 7, 199–229.
- Gu, C. (2013). *Smoothing spline ANOVA models* (Vol. 297). Springer Science & Business Media.
- Gu, C. (2014). Smoothing spline ANOVA models: R package gss. *Journal of Statistical Software*, 58, 1–25.
- Hopp, H. (2009). The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition*, 12, 463–483.
- Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29, 33–56.
- Ito, K., & Speer, S. R. (2006). Using interactive tasks to elicit natural dialogue. In S. Sudhoff, D. Lenertova, R. Meyer, P. Augurzky, S. Pappert, I. Mleinek, N. Richter, & J. Schließer (Eds.), *Language, context, and cognition: Vol. 3. Methods in empirical prosody research* (pp. 229–257). Walter de Gruyter.
- Ito, K., & Speer, S. R. (2008). Anticipatory effects 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.

- Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, 2, 647–670.
- Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32, 37–55.
- Klassen, J. (2015). *Second language acquisition of focus prosody in English and Spanish* (Unpublished doctoral dissertation). McGill University.
- Koops, C. (2010). /u/-fronting is not monolithic: Two types of fronted/u/in Houston Anglos. *University of Pennsylvania Working Papers in Linguistics*, 16, 114–122.
- Kurumada, C., Brown, M., Bibyk, S., Pontillo, D., & Tanenhaus, M. (2014). Rapid adaptation in online pragmatic interpretation of contrastive prosody. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36, 791–796.
- Kurumada, C., Brown, M., & Tanenhaus, M. (2012). Pragmatic interpretation of contrastive prosody: It looks like speech adaptation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 34, 647–652.
- Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, 98, 74–88.
- Lee, E. K., & Fraundorf, S. (2017). Effects of contrastive accents in memory for L2 discourse. *Bilingualism: Language and Cognition*, 20, 1063–1079.
- 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.
- MacDonald, M. C., Pearlmuter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703.
- McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammatical judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381–401.
- Namjoshi, J. (2015). *The processing and production of prosodic focus in French by native and non-native speakers* (Unpublished doctoral dissertation). Champaign: University of Illinois at Urbana-Champaign.
- Perdomo, M., & Kaan, E. (2019). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research*. Advance online publication. <https://doi.org/10.1177/0267658319879196>
- Reichle, R. V., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French. *Studies in Second Language Acquisition*, 36, 535–564.
- Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research*, 24, 397–430.
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., & Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. In *Proceedings of the Second International Conference on Spoken Language Processing* (pp. 867–870).
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1, 1–33.
- Takeda, A. (2018). *English prosodic marking of information structure by L1-Japanese second language learners* (Unpublished doctoral dissertation). University of Hawai‘i at Mānoa.
- The GIMP team. (2014). *GNU image manipulation program* (Version 2.8.6) [Computer software]. www.gimp.org
- Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443–467.
- Venditti, J. J., Maekawa, K., & Beckman, M. E. (2008). Prominence marking in the Japanese intonation system. In S. Miyagawa & M. Saito (Eds.), *The Oxford handbook of Japanese Linguistics* (pp. 456–512). Oxford University Press.
- Weber, A., Braun, B., & Crocker, M. W. (2006a). 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. (2006b). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, 99, B63–B72.

APPENDIX A

List of objects shown in the line drawings for each set. Grammatical gender (masculine, feminine or neuter) and corresponding definite article (*der, die, das*) in German are given in parentheses.

Set 1 (feminine; *die*)

Hand (hand)
Mütze (hat)
Puppe (doll)
Schere (scissors)

Set 3 (neuter, *das*)

Buch (book)
Eis (ice cream)
Messer (knife)
Pferd (horse)

Set 5 (masculine, *der*)

Hase (rabbit)
Schlüssel (key)
Schrank (wardrobe)
Stift (pencil)

Set 2 (masculine; *der*)

Keks (cookie)
Löffel (spoon)
Mund (mouth)
Vogel (bird)

Set 4 (feminine, *die*)

Gabel (fork)
Jacke (jacket)
Katze (cat)
Tür (door)

Set 6 (feminine, *die*)

Banane (banana)
Ente (duck)
Hose (pants)
Möhre (carrot)