Journal of Memory and Language 98 (2018) 1–11

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study

Aine Ito [a,b,*], Martin J. Pickering [a], Martin Corley [a]

[a] Department of Psychology, University of Edinburgh, United Kingdom
[b] Faculty of Linguistics, Philology & Phonetics, University of Oxford, United Kingdom

ARTICLE INFO

Article history:
Received 6 June 2016
Revision received 25 August 2017

Keywords:
Prediction
Comprehension
Bilingualism
Visual world
Phonological competitor

ABSTRACT

We report a study using the "visual-world" paradigm that investigated (1) the time-course of phonological prediction in English by native (L1) and non-native (L2) speakers whose native language was Japanese, and (2) whether the Japanese participants predicted phonological form in Japanese. Participants heard sentences which contained a highly predictable word (e.g., *cloud,* following *The tourists expected rain when the sun went behind the* ...), and viewed an array of objects containing a *target* object which corresponded to the predictable word [*cloud*; Japanese: *kumo*], an *English competitor* object whose English name was phonologically related to the predictable word [*clown*; *piero*], a *Japanese competitor* object whose Japanese name was phonologically related to the Japanese translation of the predictable word [*bear*; *kuma*], or an object that was *unrelated* to the predictable word [*globe*; *tikyuugi*]. Both L1 and L2 speakers looked predictively at the target object, but L2 speakers were slower than L1 speakers. L1 speakers looked predictively at the English competitor object, but L2 speakers did not do so predictively. Neither group looked at the Japanese competitor object more than the unrelated object. Thus, people can predict phonological information in their native language but may not do so in non-native languages.

Introduction

People predict aspects of upcoming words during language comprehension, including meaning or syntax (Altmann & Kamide, 1999; Staub & Clifton, 2006). Other studies suggest they predict phonological or orthographic word forms (DeLong, Urbach, & Kutas, 2005; Laszlo & Federmeier, 2009), but much less is known about when these predictions occur or the extent to which they depend on the availability of cognitive resources (Huettig, 2015). In this paper, we investigate the nature of phonological form prediction by tracking the eye movements of native (L1) and non-native (L2) speakers as they listen to English sentences and see pictures whose names are phonologically related to highly predictable words.

In this "visual world" paradigm, fixations to objects are driven by lexical activation (Tanenhaus, Magnuson, Dahan, & Chambers, 2000). We can therefore investigate when phonological information relevant to highly-predictable words becomes available. L2 language comprehension involves more resources than L1 language comprehension (Clahsen & Felser, 2006), and so we use a comparison of L1 and L2 comprehension to investigate whether phonological prediction is resource-intensive. Moreover, L2 comprehension is of course difficult in general, and one reason may be that L1 is not fully suppressed (e.g., Thierry & Wu, 2007). For this reason, we also tested whether L2 speakers predictively activate phonological information in their L1 by presenting them with objects whose L1 names were related to the L1 translation of the predictable word.

Prediction of phonological information in L1

Studies of word-form prediction have tended to conflate phonology with orthography, because of the close relation between the two in Western languages. For ease of exposition, we refer throughout the present paper to phonology (on the basis that the materials in our experiments are presented auditorily) but we would not be able to fully exclude an account of our evidence rooted in orthography.

Evidence about the prediction of phonological form comes exclusively from event-related potential (ERP) experiments in which participants read highly constraining sentences – that is,

* Corresponding author at: Faculty of Linguistics, Philology & Phonetics, University of Oxford, Clarendon Institute, Walton Street, Oxford OX1 2HG, United Kingdom.

E-mail address: aine.ito@ling-phil.ox.ac.uk (A. Ito).

http://dx.doi.org/10.1016/j.jml.2017.09.002
0749-596X/

sentences with a context that is very likely to be followed by a particular predictable word. There are two types of study. In the first, the predictable word is replaced by a word or nonword with a similar form to the predictable word. This stimulus elicits a smaller N400 than a word or nonword that is dissimilar to the predictable word (Ito, Corley, Pickering, Martin, & Nieuwland, 2016; Kim & Lai, 2012; Laszlo & Federmeier, 2009). In a representative study, participants read contexts such as "*The student is going to the library to borrow a*...", followed by the predictable word (*book*), an unpredictable word whose form was related to the predictable word (*hook*), or an unpredictable word whose form was unrelated to the predictable word (*sofa*). The unpredictable words *hook* and *sofa* both showed larger N400s compared to the predictable word *book*, but the N400 was reduced for the form-related word *hook* compared to the unrelated word *sofa* (Ito, Corley et al., 2016). These findings suggest that readers pre-activate the forms of predictable words. However, it is also possible that readers activated the form of the predictable word (*book*) after they encountered the form-related word (*hook*). For instance, readers might have encountered *hook* and combined it with the predictable sentence context to activate *book*.

The second type of study investigates whether form is predicted before the target stimulus is encountered. DeLong et al. (2005) found that people can predict phonological aspects of highly predictable words during reading comprehension. In their study, participants read sentence contexts that predicted a specific noun (e.g., *kite* in "*The day was breezy so the boy went outside to fly*..."). These contexts were followed by the expected noun phrase (*a kite*) or an unexpected but plausible noun phrase (*an airplane*). Unexpected nouns (*airplane*) elicited larger N400 amplitudes than expected nouns (*kite*). This N400 for expected versus unexpected nouns could indicate that participants predicted the expected noun, but could also indicate that expected nouns were easier than unexpected nouns to integrate into the context. Importantly, the authors also found a correlation between the N400 amplitudes for the preceding articles (*a*/*an*) and the cloze probabilities of these articles. The authors argued that this graded N400 for articles could not be explained by integration, and indicated that people probabilistically pre-activate an element of the phonological form of predictable words (whether it began with a vowel or a consonant).

But the reliability of this effect is under dispute. One study used the same *a*/*an* manipulation and found a larger N400 for unexpected articles compared to expected articles (Martin et al., 2013; we discuss this study in the following section *Prediction of phonological information in L2*). However, this effect of condition (expected vs. unexpected articles) was not found in DeLong et al. (2005), and Martin et al. (2013) did not report the article correlation that DeLong et al. reported. Thus, the findings from the two studies are not fully consistent. Furthermore, using materials adapted from Martin et al., another study failed to replicate Martin et al.'s effect of condition and also did not find any graded effect of article cloze probability on article N400 (Ito, Martin, & Nieuwland, 2016a). It is possible that comprehenders are not always confident that the noun (e.g., *kite*) will be the next word (e.g., the sentence could continue *an impressive kite*). But for whatever reason, it appears that N400 effects on the article do not consistently occur. It is therefore particularly important to investigate phonological prediction using another paradigm.

Even assuming prediction, a limitation of these studies is that they cannot straightforwardly reveal when the predictions occurred, because the test point occurs at, or one word before, the predictable word. So they are compatible with two accounts. On one account, comprehenders predict as soon as they are confident that the word will occur at some point downstream. In other words, they predict word form as a consequence of predicting other aspects of a word (e.g., semantics). In Ito, Corley et al. (2016), comprehenders who encountered *The student is going to the library to borrow a*..." may have predicted the form *book* after encountering *library* (or even *student*); in DeLong et al. (2005), they may have predicted *kite* after encountering *breezy day*. On the other account, they predict form immediately before the upcoming word, presumably in order to make comprehension of that word as straightforward as possible. In Ito, Corley et al., they may have predicted *book* after encountering the form-related word *hook*. In DeLong et al., they may have predicted *kite* after encountering the immediately preceding article *a*.

In a visual world experiment, eye movements are continuously recorded as participants listen to a sentence. If the scene contains an object whose name is related in form to the predictable word (e.g., a hook in Ito, Corley et al., 2016), then participants who preferentially look at that object must have predicted the form of the predictable word (because the form-related word is not related to the predictable word or to the context in any other way). These prediction-driven fixations may therefore occur much earlier than the predictable word. Thus, we expected that our study would provide more information about the time-course of prediction than previous ERP studies.

Our experimental logic is based on one used by Rommers, Meyer, Praamstra, and Huettig (2013) who investigated the prediction of physical aspects (shape) of the referents of upcoming words. Their participants heard highly constraining sentences (e.g., "*In 1969 Neil Armstrong was the first man to set foot on the moon*") while viewing a scene containing a picture representing the predictable target object (the moon), an object of a similar shape to the target object (a tomato), or an unrelated object (rice). The scenes also contained three unrelated distractor objects. If participants pre-activated the shape of the target word, they would be expected to fixate the similar-shaped object, as a result of their shape-related similarity (competitor effect). Participants fixated the similar-shaped object more than the unrelated objects before the target word could be processed (assuming a 200 ms delay to initiate eye movements; Saslow, 1967). Thus, these findings support pre-activation of shape information.

The present study was closely modelled on the design used by Rommers et al. (2013). To investigate pre-activation of phonological information, we used phonologically related, rather than shape-related, competitors. We did not present a predictable target object when its competitor object was present. The primary advantage of this design is that it should prevent looks to the competitor object being swamped by looks to the target object. In other words, the absence of the predictable object should give participants more opportunity to fixate on the competitor object.

Prediction of phonological information in L2

The resources available to L2 speakers are more limited than the resources available to L1 speakers. Compared to L1 speakers, L2 speakers may be slower to access lexical information or have weaker semantic networks (e.g., Ivanova & Costa, 2008). They may also be less good at using syntactic information (Clahsen & Felser, 2006), or may comprehend less automatically (Segalowitz & Hulstijn, 2009). Thus, we expected that L2 speakers would predict to a lesser extent than L1 speakers.

There is evidence that L2 speakers can predict some features of upcoming words, including semantic (Chambers & Cooke, 2009; Ito, Corley, & Pickering, 2017) or syntactic information (Foucart, Martin, Moreno, & Costa, 2014; Foucart, Ruiz-Tada, & Costa, 2016). However, it is less clear whether L2 speakers predict phonological information. As we have noted, Martin et al. (2013) used DeLong et al.'s (2005) paradigm, and found that L1 speakers showed a larger N400 for pre-nominal articles that were incompatible with the predictable word compared to articles that were compatible with the predictable word (e.g., *an* vs. *a*; in the context where *kite* was expected). But they did not find this effect in L2 speakers. Although L2 speakers, like their L1 counterparts, showed an N400 effect for unpredictable nouns relative to predictable nouns, their N400 responses did not differ at the preceding articles (even though they were familiar with the *a*/*an* rule in English). The results suggest that L2 speakers do not predict phonological information like L1 speakers (see also Ito, Martin, & Nieuwland, 2016b).

It is therefore possible that L2 speakers rarely or never predict phonological information. However, the cloze probabilities in Martin et al. (2013) were not particularly high (69% in L1 speakers and 65% in L2 speakers). In general, cloze probability is likely to be lower for L2 speakers than for L1 speakers, presumably because L2 speakers have had less exposure to the language. Thus, L2 speakers may be less able to use contextual information for prediction compared to L1 speakers. However, it is possible that L2 speakers do in fact predict phonological information when the relevant word is highly predictable. We address this issue by using highly predictable target words.

Another potential influence on prediction in L2 comprehension is competing information from L1, if this is automatically activated. If so, predictions made in L1 during L2 comprehension might interfere with predictions made in L2. In other words, an apparent lack of prediction in L2 might be found, not because L2 speakers are not predicting, but because they are attempting to predict in more than one language simultaneously. In fact, there is some evidence that L2 speakers activate phonological information in L1 during L2 comprehension. For example, Mishra and Singh (2014) found that Hindi–English bilinguals activated the L1 (Hindi) translation-equivalent word form of an L2 (English) target word after hearing that target word in a sentential context (see Thierry & Wu, 2007, for ERP evidence using a different paradigm). Other studies report cognate facilitation effects, in which participants process L2 words that share phonological or orthographic forms with an L1 word faster than words that do not (De Groot & Nas, 1991; Dijkstra, Van Jaarsveld, & Brinke, 1998; Libben & Titone, 2009). These results suggest that L2 speakers may activate translation equivalents (including their phonology) in their L1 during L2 comprehension.

But do people who are comprehending in their L2 *pre*-activate phonological forms in their L1? One reason why pre-activation might occur is if the evidence of L1 phonological activation during L2 comprehension is part of a general tendency for activation to propagate from (pre-)activated semantics to L1 phonology in L2 contexts. To date, there is no evidence for L1 phonological pre-activation during L2 listening comprehension. The current study therefore investigated whether L1 Japanese – L2 English speakers who are listening to English sentences pre-activate Japanese phonological forms for words that are the translation equivalents of highly predictable English words.

Phonological prediction in L2 (like phonological prediction in L1) has only been tested in ERP paradigms. Here, we test L1 and L2 prediction in the visual world paradigm, allowing us to investigate the time-course of phonological prediction in proficient and less-proficient speakers, and to answer the questions of when prediction occurs and whether it requires cognitive resources.

The current study

Our study investigated the pre-activation of phonological information when a specific word was highly predictable, by L1 English speakers (who reported no knowledge of Japanese), and L2 English speakers whose native language was Japanese. Considering the possibility that phonological prediction is not strong, we constructed very high-cloze sentences in order to maximize the likelihood of detecting any effect. As the evidence for phonological prediction before the target noun (e.g., ERP experiments using *a*/*an*) is equivocal, we used an eye-tracking paradigm, which has not been used to investigate phonological prediction. This also allowed us to investigate the time-course of phonological predictions.

Participants listened to sentences which contained a highly predictable word, while viewing a scene containing one of four critical objects: a *target* object whose English name corresponded to the predictable word [*cloud*; Japanese: *kumo*], an *English competitor* object whose English name was phonologically related to the predictable word [*clown*; *piero*], a *Japanese competitor* object whose Japanese name was phonologically related to the Japanese translation of the predictable word [*bear*; *kuma*], or an object that was *unrelated* to the predictable word [*globe*; *tikyuugi*]. They also saw three unrelated distractor objects.

If participants predict highly predictable target words, they should fixate more on the target objects than on unrelated objects before hearing the target word. Such predictive looks would not demonstrate that participants predict phonological information, since this effect could occur as long as participants predict some information about target words (e.g., meaning). The critical hypotheses concern the English competitor condition and the Japanese competitor condition. If participants predict phonological information, they should fixate on objects corresponding to English competitors more than unrelated objects. If L2 speakers pre-activate phonological information about the Japanese translation of the predictable target word, they should fixate on objects corresponding to Japanese competitors more than unrelated objects.

Methods

Participants

Twenty-four native English speakers who reported no knowledge of Japanese and 24 L2 English speakers whose L1 was Japanese participated in the experiment. Four further participants (two L1 participants and two L2 participants) were excluded from the analyses because they almost never (less than 3% of the time) fixated the depicted objects (experimental items and filler objects); cf. Hintz and Meyer (2015). All participants had normal vision and reported no language disorders. All participants were resident in Edinburgh at the time of testing.

L2 participants filled in a language background questionnaire before the experiment. Their mean age of first exposure to English was 10 years (range = 5–15 years). The mean length of exposure to English (the total length of any form of regular exposure to English, including both classroom and non-classroom situations) was 13 years (range = 4–20 years). IELTS scores (www.ielts.org) were reported by 15 participants (*Mean score* = 7, range = 6.5–8).[1] L2 participants also self-rated their English proficiency in reading, writing, speaking, and listening separately on a scale from 1 (not good at all) to 10 (very good). The mean self-rated proficiency averaged over the four measures was 7.0 (*SD* = 1.5). This score was similar to the score Martin et al. (2013) reported for their L2 participants (*M* = 7.6, *SD* = 1.0, on the same 1–10 scale).

Stimuli

Experimental stimuli consisted of 16 sentences, each paired with one of four visual scenes (see Appendix for the full set of

[1] IELTS is an English proficiency test for people who want to study or work in an English-speaking country. The score is assessed on a nine-band scale from non-user (band score 1) through to expert (band score 9).

items). The experimental sentences each contained a highly predictable target word (e.g., *cloud* in "*The tourists expected rain when the sun went behind the cloud, but the weather got better later.*") at different positions in the sentence (range = 9th-20th word, $M = 13.7$, $SD = 2.6$) but never sentence-finally. The sentences consisted of a mean of 17.6 words ($SD = 1.4$, range = 16–21 words). There were an additional 16 filler sentences, of similar length to the experimental sentences. The sentences were recorded by a male native British English speaker, and sampled at 48 kHz with a format of 32-bit float. The speaker read the sentences at a rate of approximately 2.6 syllables per second with some space between phrases. The mean duration of experimental sentences was 10.1 s.

The predictability of the target words was assessed using a cloze probability test. Twelve native English speakers and 12 Japanese-English late bilinguals who were studying in the UK read sentences truncated before the target word, and completed each sentence fragment using the first word that came to mind. None of them participated in the eye-tracking experiment. The mean cloze probability of the target word was 97.5% ($SD = 3.7$, range = 91.7–100%) in L1 speakers and 88.6% ($SD = 7.1$, range = 81.8–100%) in L2 speakers. The L2 cloze probability was significantly lower than the L1 cloze probability, $t(30) = 3.9$, $p < 0.001$, but was considerably higher than the L2 cloze probability reported in Martin et al. (2013) ($M = 65\%$, $SD = 26$). We expected that the high cloze probabilities would maximize the likelihood of detecting effects of phonological prediction and in particular its time course.

Each of the visual scenes contained four objects: a critical object and three distractors. In the target condition, the critical object corresponded to the target word (e.g., *cloud* [Japanese: *kumo*]). In the English competitor condition, the English name of the critical object phonologically overlapped at onset with the target word (e.g., *clown* [*piero*]). In the Japanese competitor condition, the Japanese name of the critical object phonologically overlapped at onset with the Japanese translation of the target word (e.g., *bear* [*kuma*]). The mean number of phonemes shared between target words and English competitor words was 2.9 ($SD = 0.83$) out of a mean of 4.4 phonemes (66.2%), and that between Japanese translations of target words and Japanese competitor words was 2.6 ($SD = 0.60$) out of a mean of 4.9 phonemes (53.8%). English names and Japanese names of the Japanese competitor objects were both unrelated to any of the English names of the target, English competitor, and unrelated objects. English and Japanese names of each critical object were also unrelated to each other. In the unrelated condition, the name of the critical object did not have phonological onset overlap with the predictable word or its Japanese translation (e.g., *globe* [*tikyuugi*]). We created the unrelated condition as a baseline, so that we could keep the distractors the same across conditions. All four objects were semantically unrelated to each other.

We conducted a picture naming test to assess name agreement for the depicted objects. Native English speakers and Japanese-English late bilinguals who did not participate in the eye-tracking experiment saw pictures of objects and gave the first word that came to mind when they saw each picture. Some of the items were changed and re-tested, and every picture in the final set of stimuli was tested by at least 12 participants from each group. The naming agreement for objects, for L1 and L2 speakers respectively, was L1: 94.2% ($SD = 6.3$, range = 83.3–100%) and L2: 93.2% ($SD = 8.4$, range = 76.9–100%) in the target condition, L1: 86.6% ($SD = 13.2$, range = 61.1–100%) and L2: 86.7% ($SD = 10.8$, range = 66.7–100%) in the English competitor condition, L1: 92.8% ($SD = 9.4$, range = 66.7–100%) and L2: 93.8% ($SD = 9.9$, range = 75.0–100%) in the Japanese competitor condition, and L1: 92.2% ($SD = 8.4$, range = 75–100%) and L2: 94.4% ($SD = 9.9$, range = 66.7–100%) in the unrelated condition. Another group of Japanese-English late bilinguals from a similar population completed the Japanese version of the same pre-test. In this Japanese naming pre-test, the instructions were translated into Japanese, and English was not used throughout the test. The naming agreement for critical objects was 91.7% ($SD = 12.5$, range = 58.8–100%) in the target condition, 87.9% ($SD = 16.6$, range = 41.2–100%) in the English competitor condition, 90.1% ($SD = 13.7$, range = 64.7–100%) in the Japanese competitor condition, and 97.4% ($SD = 4.3$, range = 88.2–100%) in the unrelated condition.

All of the visual stimuli were shown twice, once in an experimental trial and once in a filler trial. Each experimental list comprised two half-lists, each made up of the 16 visual stimuli paired with 8 experimental and 8 filler recordings. Matched visual stimuli contained the same objects, but the quadrants in which these objects appeared were varied. Visual stimuli which were paired with experimental items in one half-list were paired with fillers in the other half-list, and vice versa. Experimental pictures were counterbalanced in the full lists, resulting in 4 different sets of items, or 8 experimental lists in total.

Critical objects appeared at each of the four quadrants equally frequently. Filler sentences mentioned one of the three distractor objects in the visual scene 75% of the time, so together with the experimental sentences (which mentioned one of the four objects 25% of the time), 50% of sentences referred to an object in the visual scene. An example item with four conditions is shown in Fig. 1.

Procedure

All participants were tested in the University of Edinburgh. The experiment started with a picture familiarisation task. First, participants saw the 64 experimental objects one by one with their English name presented both visually and auditorily at the same time (the names were recorded by the same speaker in the same way as the experimental sentences). Participants were instructed to associate the images with the words, so that they would be able to name them later. After that, they were asked to name each object using the word given earlier. Incorrectly named objects (0.9% in L1 participants, 6.7% in L2 participants) were repeated until participants named them correctly. The number of incorrectly answered objects in each critical condition was 5 (target), 5 (English competitor), 8 (Japanese competitor), and 4 (unrelated).

In the eye-tracking experiment, participants were seated in front of a computer screen, and they were asked to listen to the sentences and judge whether each sentence mentioned any of the objects in the display. Participants' eye movements were recorded using an EyeLink 1000 Tower mount eye-tracker sampling at 500 Hz. After the instructions, each participant placed their chin on a chin rest, and the eye-tracker was calibrated using the nine-point calibration grid. The experiment started with two practice trials, after which participants were given a chance to ask questions. The pictures were presented on a viewing monitor at a resolution of 1024×768 pixels. Each trial started with a drift correction, which was followed by a 500 ms blank screen. The visual scene was presented 1000 ms before the onset of predictable words in experimental trials. On filler trials, the presentation was 1000 ms before the onset of a word that referred to a distractor or at an arbitrarily chosen mid-sentence position when the sentence did not mention anything in the scene. The picture stayed on the screen for 750 ms after the offset of the spoken sentence. The picture disappeared and the question "*Did the sentence mention any of the pictures?*" appeared immediately. The participant responded by pressing "1" for "Yes" and "2" for "No" on the keyboard, and the next trial began immediately. No feedback was given during the experiment. The session took about 30 min.

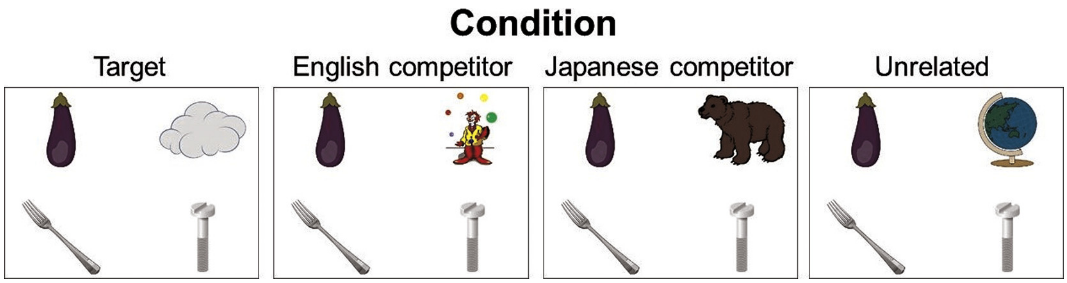

Fig. 1. Example visual scene in four conditions for the experimental sentence "*The tourists expected rain when the sun went behind the* ***cloud****, but the weather got better later.*" The object depicted at the top right corner is the critical object for this item. The visual stimuli were also paired with the filler sentence "*The waiter immediately came over to the table when the woman carelessly dropped her* ***fork***."

Results

Picture naming accuracy

We calculated the proportion of trials where the first name participants produced was correct. The mean picture naming accuracy was 99.1% (*SD* = 1.3%) for L1 speakers and 93.3% (*SD* = 7.1%) for L2 speakers. The high accuracy suggests that the pictures were relatively easy to associate with the intended names.

Comprehension question accuracy

The mean accuracy for the comprehension questions in the experimental trials was 100% for L1 speakers and 99.2% (*SD* = 2.1%) for L2 speakers. Incorrectly answered trials were excluded from the eye-tracking analyses.

Eye-tracking data analyses

Two items were excluded from the eye-tracking analyses for L1 speakers because the English competitor object in these items attracted significantly more looks than the unrelated object within 1000 ms after the picture onset when the pictures were presented with a neutral sentence that was unrelated to the English competitor objects in filler trials (see the analysis for filler trials in the section below). This left 14 items for the analyses.[2] L2 speakers showed no such preference to fixate any given item over another, so no items were excluded.

We first analysed the eye-tracking data separately for L1 and L2 speakers using a linear mixed-effects model, with the lme4 package (Bates, Maechler, & Dai, 2008) in R (R Development Core Team, 2015). The proportions of time spent fixating on target, English competitor, Japanese competitor, and unrelated objects were calculated separately for each 50 ms bin relative to the target noun onset (following Altmann & Kamide, 1999). Blinks and fixations outside the computer screen were coded as 0 (i.e., no fixation on any of the objects) and were included in the data. To explore the time-course of effects, we ran the model for every 50 ms bin from 1000 ms before target word onset to 1000 ms after the onset. The model evaluated the arcsine-transformed fixation proportions on critical objects as predicted by condition, for each bin. Condition was dummy-coded, so that we could test effects of each critical condition relative to the unrelated baseline condition (target vs. unrelated, English competitor vs. unrelated, and Japanese competitor vs. unrelated). The model included random intercepts by participants and by items (Barr, 2008). Random slopes were not included because the models with them did not converge for several of the bins. To account for familywise error, our conclusions are based on periods where a series of consecutive bins showed a significant difference (cf. Borovsky, Elman, & Fernald, 2012). To confirm these conclusions, we additionally conducted a growth curve analysis that did not involve multiple comparisons (Mirman, Dixon, & Magnuson, 2008).

The growth curve analysis included a third-order (cubic) orthogonal polynomial to capture non-linear changes in fixation proportions over the time-course. This analysis tested for effects of time, and an interaction of condition by language group, while avoiding multiple comparisons. This analysis did not include the Japanese competitor condition because we found no Japanese competitor effects in the linear mixed-effects model analysis. For the growth curve analysis, we coded fixations binomially depending on whether the relevant object was fixated or not during each 50 ms bin. We then transformed the coded data into log odds (suitable to test effects on a categorical variable). We constructed a model evaluating the transformed fixation proportions predicted by fixed effects of condition (target vs. unrelated and English competitor vs. unrelated) and language group (L1 vs. L2), and the interaction of the two with all time terms. Condition and language group were both dummy-coded. In the by-participant analysis, the model additionally included participant random effects on all time terms and participant-by-condition random effects on all time terms except the cubic (estimating random effects is "expensive" in terms of the number of observations required, so this cubic term was excluded because it tends to capture less-relevant effects occurring at the end of the analysis time window). In the by-item analysis, the model included item random effects and item-by-condition random effects on all time terms except the cubic, instead of participant- or participant-by-condition random effects. In both the linear mixed-effects model and growth-curve analyses, significance of effects was determined by assessing whether the associated *t*-statistics had absolute values which exceeded 2 (Baayen, Davidson, & Bates, 2008).

Eye-tracking data

Fig. 2 shows the proportions of fixations on target, English competitor, Japanese competitor, and unrelated objects for L1 speakers and for L2 speakers. The figure suggests that both L1 and L2 speakers fixated target objects more than unrelated objects before they heard the target word, but the time-course of fixating English competitor objects over unrelated objects was different in the two groups. The binned linear mixed-effects model for L1 speakers showed that L1 speakers were more likely to fixate target objects than unrelated objects from 600 ms before the acoustic onset of the target word until 1000 ms after the target word onset. This bias towards the target objects indicates participants' sensitivity to the target word predictability. Critically, L1 speakers were also more likely to fixate English competitor objects than unrelated objects

[2] We also analysed the data including all the 16 items. A linear-mixed model including the same predictors showed the same pattern of effects as the main analyses.

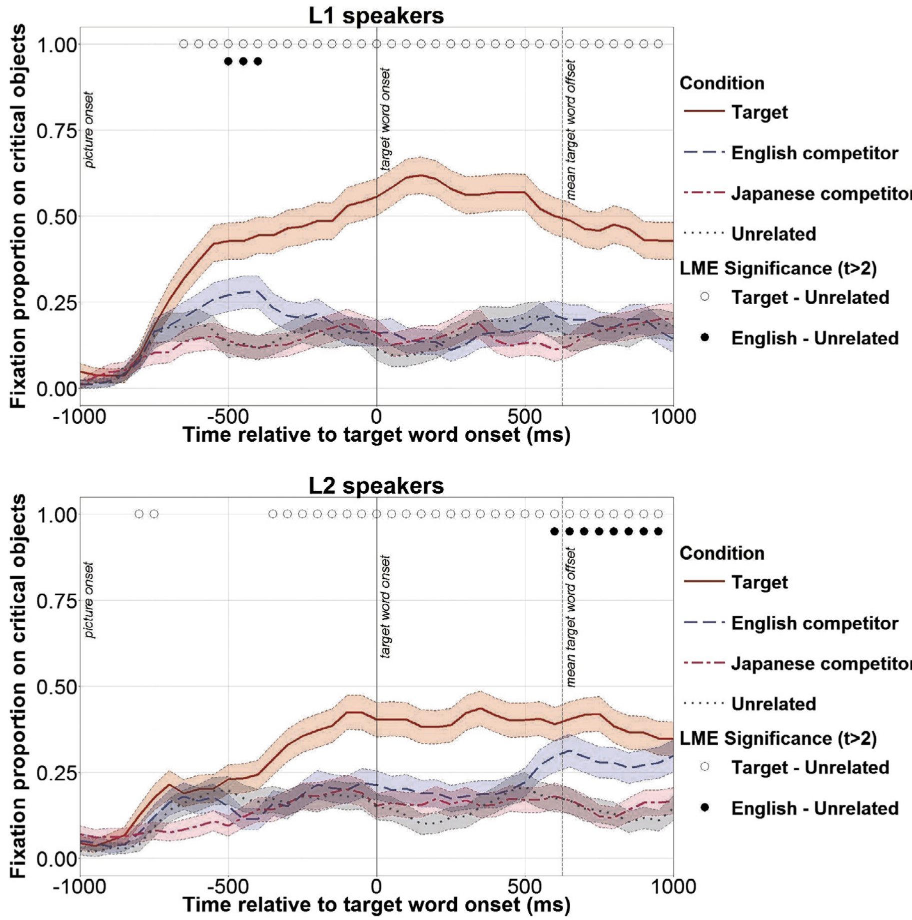

Fig. 2. Eye-tracking results in L1 speakers (top) and in L2 speakers (bottom). Time-course graph showing fixation proportion on target, English competitor, Japanese competitor, and unrelated objects. Time 0 ms shows target word onset. The dashed vertical line (y = 625 ms) indicates the mean target word offset. Circles at the top of the graph show significant differences ($|t| > 2$) between the target and unrelated conditions (open circle, ○), and between the English competitor and unrelated conditions (solid circle, ●), corresponding to the time on the x-axis. Transparent thick lines are error bars representing standard errors.

between 500 ms before the target word onset and 350 ms before the target word onset. As predicted, L1 speakers did not show any bias towards Japanese competitor objects relative to unrelated objects in any of the time windows. These results suggest that L1 speakers predicted target words and pre-activated their phonological information.

Equivalent analyses for L2 speakers showed that L2 speakers were more likely to fixate target objects than unrelated objects between 800 ms and 700 ms before the target word onset, and from 350 ms before the target word onset until 1000 ms after the target word onset. This finding suggests that L2 speakers also predicted some information about target words. However, L2 speakers fixated English competitor objects more than unrelated objects only from 600 ms until 1000 ms after the target word onset. This late effect of the English phonological competitor suggests that L2 speakers did not predict phonological information, but that they did activate phonological information associated with the target word after encountering it (e.g., via priming). At no point from 1000 ms before the target word onset to 1000 ms after the target word onset were there any differences between the Japanese competitor condition and the unrelated condition. Thus, there was no evidence that L2 speakers ever activated the phonology of the Japanese translations of predictable words.

To test whether L1 speakers and L2 speakers differed in predictive eye movements, we used a growth curve analysis and tested for an interaction of language nativeness with the target prediction effect and English phonological competitor effect during the prediction time window (i.e., the 1000 ms window from picture onset to target word onset). The by-participant analysis model revealed significant effects of target condition, $\beta = 1.4$, $SE = 0.31$, $t = 4.7$, and English competitor condition, $\beta = 0.62$, $SE = 0.31$, $t = -2.0$, relative to the unrelated condition (baseline) on the intercept term, indicating that there were more looks to target objects and English competitor objects relative to unrelated objects overall. The interaction of target condition (target vs. unrelated) by language group was significant on the cubic term, $\beta = -1.6$, $SE = 0.60$, $t = 2.6$, and the interaction of English competitor condition (English competitor vs. unrelated) condition by language group was significant on the quadratic term, $\beta = -3.1$, $SE = 1.4$, $t = -2.2$, and on the cubic term, $\beta = 1.7$, $SE = 0.60$, $t = -2.9$. An inspection of Fig. 3 suggests that the target condition (target vs. unrelated) by language group interaction was because the fixation proportion difference between the target and unrelated conditions reached a peak and became stable earlier in L1 speakers than in L2 speakers. The English competitor condition (English competitor vs. unrelated) by language group interaction occurred because the fixation difference between the English competitor and unrelated conditions increased over time and decreased after reaching a clear peak in L1 speakers, whereas the conditions did not differ throughout the entire time window in L2 speakers (Fig. 3, top). The by-item analysis model also revealed a significant effect of target condition on the intercept term, $\beta = 1.9$, $SE = 0.33$, $t = 5.7$. The effect of English competitor approached significance, $\beta = 0.65$, $SE = 0.33$, $t = 2.0$, on the intercept term. In the comparison of target and unrelated con-

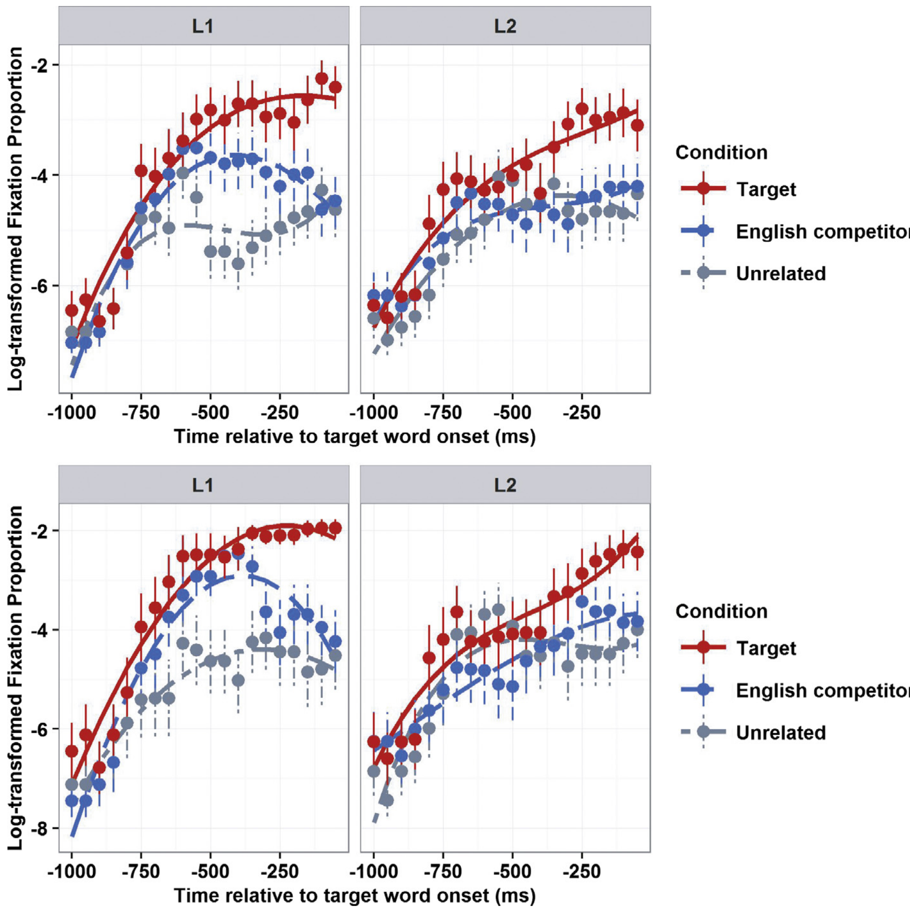

Fig. 3. Growth curve analysis model fits (lines) of the fixation data in the target, English competitor and unrelated conditions in the L1 group (left) and the L2 group (right). The top graph shows the by-participant analysis and the bottom graph shows the by-item analysis. Error bars represent ±1 *SE*.

ditions, the interaction with language group was significant on the intercept term, $\beta = -1.1$, $SE = 0.21$, $t = -5.3$, and on the quadratic term, $\beta = 2.7$, $SE = 0.92$, $t = 3.0$. The interaction on the quadratic term captured the pattern wherein L1 speakers showed the largest difference between the conditions approximately in the middle of the time window, but L2 speakers showed the largest difference towards the end of the time window. In comparison of English competitor and unrelated conditions, the interaction with language group was also significant on the quadratic term, $\beta = 4.1$, $SE = 0.92$, $t = 4.5$, suggesting that the difference between the English competitor and unrelated conditions reached a peak approximately in the middle of the time window in L1 speakers, whereas the difference for L2 speakers appears to be smallest in the corresponding time window (Fig. 3, bottom).

Since the by-participant and by-item analyses showed slightly different patterns, Fig. 3 presents a graph for each analysis. Despite the different patterns in the time-course, interactions of condition by language group were found in both analyses, and showed that predictive looks to target objects occurred later in L2 speakers than in L1 speakers, and that the English competitor effect in the prediction time window was evident in L1 speakers but not in L2 speakers.

We also analysed the filler trials in order to examine whether there was any visual bias towards critical objects irrespective of the predictive contexts. As reported above, two items were excluded from the analysis for L1 speakers after an initial analysis of the filler items. In the remaining 14 items, the linear mixed-effects model for L1 speakers did not show any fixation proportion differences between conditions ($|t|\text{s} < 2$), except that Japanese competitor objects attracted more fixations than unrelated objects in a single 550–600 ms time bin after the picture onset, and that English competitor objects attracted more fixations than unrelated objects from 850 ms to 1000 ms after the picture onset ($|t|\text{s} > 2$). Since these effects do not pattern with the data in experimental trials, the predictive English phonological competitor effect obtained in experimental trials cannot be explained by any visual biases towards the competitor objects. In L2 speakers, fixation proportions did not differ between any of the condition pairs. Therefore, the English competitor effect in the late time window in L2 speakers cannot be attributed to visual attractiveness of the competitor objects.

We further explored the relationship between the English competitor effects in L2 speakers and their length of exposure to English. For each L2 participant, we calculated a difference in the mean arcsine-transformed fixation proportions between the English competitor and unrelated conditions in a time window from 600 ms to 1000 ms after the target word onset (we chose this time window because the difference between the English competitor and unrelated conditions was not significant from 1000 ms onwards).[3] We used this as a measure of the English competitor effect, and computed a correlation with the L2 participants' self-reported length of exposure to English. As shown in Fig. 4, we found a positive correlation between the two measures, $r(22) = 0.55$, $p < 0.01$; L2 speakers who had been exposed to English for longer showed a stronger English phonological competitor effect.

[3] We also inspected the eye-tracking data in highly proficient L2 participants (median-split according to the length of exposure to English, $M = 16$ years), but there was no hint of evidence for an English phonological competitor effect before the target word onset.

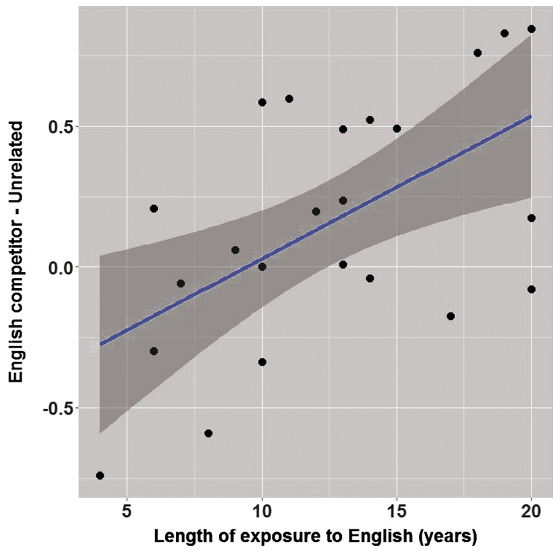

Fig. 4. Correlations between the arcsine-transformed fixation proportion difference in the English competitor condition and in the unrelated condition (in the 600–1000 ms window) and the length of exposure to English in L2 speakers.

Discussion

We investigated the time-course of prediction of the phonological information associated with highly predictable words in L1 and L2 speakers. Both groups of participants showed increased looks to target objects well before they were mentioned, although these predictive looks occurred later in L2 speakers. L1 speakers were more likely to fixate objects whose English name was phonologically related to the predictable word relative to objects whose English name was phonologically unrelated to the predictable word from 500 ms before the predictable word onset (hereafter, *phonological competitor prediction effect*). However, L2 speakers did not show such a tendency until 600 ms after the predictable word onset (hereafter, *phonological competitor priming effect*). Their tendency to fixate phonologically related objects over unrelated objects positively correlated with their length of exposure to English. L2 speakers did not fixate objects whose name was phonologically related to the Japanese translation of the predictable word.

The timing of phonological prediction in L1

The phonological competitor prediction effect patterns with the data reported in Rommers et al. (2013), who found that shape competitor objects of predictable words attracted more looks than unrelated objects prior to the mention of predictable words. In both studies, effects started to emerge about 500 ms after the objects appeared. But interestingly, the shape competitor effect in Rommers et al. (2013) lasted for much longer (about 1000 ms duration, based on visual inspection) than the phonological competitor prediction effect in our study (about 150 ms duration).

The difference in the time-course between phonological and shape predictions may be because shape relates to semantics, and semantic competitor effects are generally stronger and more sustained than phonological competitor effects (Hintz & Huettig, 2015; Huettig & McQueen, 2007). Semantic prediction could entail broader and less detailed information (e.g., a semantic category such as fruits) than phonological prediction, and the predicted semantic information could be relevant to the context for longer than the predicted phonological information. Thus, people may predict semantic information more confidently or more strongly. Alternatively (or additionally), the stronger competitor effects for shape information may be due to the visual world setting, in which the task and the dependent measure (eye movements) are heavily visual (De Groot, Huettig, & Olivers, 2016). In this setting, retrieval of shape information from a visual scene is likely to be easier than retrieval of phonological information. Of course, this conclusion about stronger shape prediction compared to phonological prediction is based on a comparison between experiments conducted in different laboratories under different conditions, so it is quite speculative (we note for example that items were previewed for 1000 ms in the present study compared to 500 ms for Rommers et al., 2013).

Another way in which the present study differed from Rommers et al. (2013) was that we used a picture familiarisation task while Rommers et al. did not. The aim of this task was to ensure that every participant associated each picture with the same word (cf. Allopenna, Magnuson, & Tanenhaus, 1998), but it might have made participants more engaged in phonological processing during the eye-tracking task. It might be possible to address this issue by using printed words instead of pictures as visual stimuli (McQueen & Viebahn, 2007), where each entity in the display unambiguously represents a word.

The phonological competitor prediction effect is consistent with ERP evidence for pre-activation of word form during L1 reading, assuming that the ERP findings demonstrate word form prediction (Ito, Corley et al., 2016; Kim & Lai, 2012; Laszlo & Federmeier, 2009). Importantly, the effect in the present study emerged 500 ms before the onset of predictable words, suggesting that participants had predicted phonological information of the predictable word by this point. Thus, they did not simply predict the form of the immediately following word (as may have been the case in DeLong et al., 2005). Instead, they predicted the forms of words that were predictable and likely to occur downstream.

No evidence for phonological prediction in L2

L2 speakers made predictive eye movements to target objects, but their predictive eye movements were slower than those of L1 speakers. In addition, there was no indication that L2 speakers predicted phonological information. This latter finding is consistent with Martin et al. (2013), even though word predictability was higher in our study (89%) than in Martin et al. (65%). It fits with the proposal that language understanding does not require prediction at all levels of representation (Huettig & Mani, 2016). This finding may also suggest that L2 speakers do not have the resources to predict detailed information about predictable words. We know that successful prediction can facilitate comprehension (Frisson, Rayner, & Pickering, 2005; Rayner, Slattery, Drieghe, & Liversedge, 2011; Smith & Levy, 2013), and these limitations may underlie some of the difficulty of L2 comprehension.

A possible account of these findings can be made in terms of prediction-by-production (e.g., Pickering & Garrod, 2007; Pickering & Garrod, 2013), in which comprehenders predict using the mechanisms involved in language production. More specifically, they covertly imitate the language they encounter, and then derive the intention (or production command) that underlies language production. They then use the processes involved in language production to predict the upcoming utterance. When predicting words, comprehenders therefore pre-activate the representations (e.g., meaning, sound) associated with a predictable word, just as they would when producing that word. One way they can do this is using what we term *prediction-with-implementation*, whereby prediction involves going through the representational stages in order, for example computing semantics before phonology (see Ito, Corley et al., 2016). In addition, since full implementation of the production system requires time and resources, phonological predictions are less likely to occur when resources are limited.[4] The additional difficulty associated with L2 production means that phonological representations are less likely to be constructed than is the case in L1 production.

Although prediction-by-production plausibly accounts for our findings, the differences between L1 and L2 prediction might also have occurred because the cloze probabilities for our sentences were higher for L1 speakers (98%) than for L2 speakers (89%). In other words, target words were somewhat more predictable for L1 speakers than for L2 speakers. As we noted earlier, however, the L2 cloze probability in the current study was extremely high in comparison to other studies on prediction during L2 comprehension (81% in Foucart et al., 2014; 61% in Ito, Martin et al., 2016b; 65% in Martin et al., 2013), and almost equivalent to the cloze probabilities in studies that have found evidence for phonological or orthographic prediction in L1 speakers (e.g., 90% in Kim & Lai, 2012; 89% in Laszlo & Federmeier, 2009).

Another possibility is that associations between the pictures and their names in L2 speakers were not as strongly established as in L1 speakers, and so the pictures might have triggered phonological activation to a lesser extent in L2 speakers. But this explanation is unlikely because all participants successfully completed picture familiarisation. The lack of evidence for L2 phonological prediction cannot be due to L2 speakers' insensitivity to the phonological overlap between target words and their English competitor words, because we found an English phonological competitor effect in L2 speakers (i.e., they tended to fixate English competitor objects over unrelated objects). However, because this phonological competitor effect did not manifest until well after the predictable word onset, it is unlikely to reflect phonological prediction, but it suggests that hearing a target word (e.g., *cloud*) leads to spreading activation to phonologically related words (e.g., *clown*). This phonological competitor priming effect correlated with the length of exposure to English of the L2 speakers. This finding fits with studies that have found a stronger phonological competitor effect for more proficient L2 speakers relative to less proficient L2 speakers (Blumenfeld & Marian, 2007; 2013).

No evidence for L1 activation during L2 comprehension

We found no evidence that L2 speakers activated L1 translations of the English target words. Thus, the results do not support the proposal that the lack of phonological prediction in L2 is due to an interference from phonological pre-activation in L1. Since our picture naming pre-test on L2 speakers found similarly high name agreement for English and Japanese, the lack of evidence for Japanese activation cannot be explained by a difference in name agreement.

It is possible that Japanese-English bilinguals did not activate Japanese at all during the experiment, because all the experimental setting was in English. However, this explanation is not likely given the evidence that L2 speakers activate their L1 even when L1 is irrelevant (Thierry & Wu, 2007; Wu, Cristino, Leek, & Thierry, 2013). Alternatively, a Japanese phonological competitor effect may have been too weak to affect eye movements – a possibility that seems reasonable given the relative weakness of the English phonological competitor effect (about 10% difference in fixation proportion, for about 150 ms). It is also possible that no Japanese competitor effect occurred because the picture naming task was conducted in English but not in Japanese before the main experiment; we did this in order to keep the experiment comparable between L1 and L2 speakers. Picture naming might have boosted word form activation or facilitated retrieval of picture names, for example via lexical priming from production to comprehension (Wheeldon & Monsell, 1992), and so the activation of English but not Japanese phonology may have been enhanced. Finally, it is conceivable that the lack of Japanese competitor effect might have been due to the lack of orthographic overlap. English target words and their phonological competitors were also orthographically related, whereas the Japanese translations and their Japanese competitors were phonologically related (e.g., *kumo* – *kuma*) but not orthographically related (e.g., 雲 – 熊).

Conclusion

Our visual world study found that both L1 and L2 speakers made predictions about upcoming words. However, L2 speakers made predictive looks later than L1 speakers, and we have argued that resource limitations associated with L2 processing delayed predictive processing. L1 speakers appear to predict specific phonological information associated with highly predictable words, but L2 speakers do not. This evidence suggests a limitation to phonological prediction, and is compatible with the suggestion that phonological prediction may not always occur.

Acknowledgments

We thank Dan Mirman and Ian Finlayson for advice with the growth curve analysis, Patrick Sturt for help with the stimuli recording, and Max S Dunn III for help with the data collection.

Appendix A.

Critical sentences and the names of critical objects for each condition (predictable, English competitor, Japanese competitor, unrelated, respectively). Predictable words are underlined. The names in square brackets are the Japanese translations written in alphabets. Cloze values collected from L1 and L2 speakers are shown after each sentence in brackets respectively. Items No. 8 and 10 were removed from the analyses in L1 speakers.

No.	Sentence (L1 cloze; L2 cloze)	Object names
1	In order to have a closer look, the dentist asked the man to open his <u>mouth</u> a little wider. (100; 100)	mouth[kuti]/ mouse [nezumi]/ socks[kutusita]/ bone[hone]
2	It takes about an hour to fly from Edinburgh to London, and about 4 hours by <u>train</u>, usually. (92; 83)	train[densya]/ tray[tore-]/ calculator[dentaku]/ goat [yagi]
3	In an emergency, we cannot use a lift; instead, we need to use the <u>stairs</u> for our safety. (100; 83)	stairs[kaidan]/ stapler [hottikisu]/ seashell [kaigara]/ onion [tamanegi]
4	If the sun comes out during a heavy shower, you can sometimes see a <u>rainbow</u> in the sky. (100; 92)	rainbow[niji]/ radio [rajikase]/ meat[niku]/ barrel[taru]
5	The tourists expected rain	cloud[kumo]/ clown

(*continued on next page*)

[4] Pickering and Garrod (2013) also proposed another form of prediction-by-production called *prediction-by-simulation*, under which people use their forward models to predict their own utterances. Critically, unlike *prediction-by-implementation*, prediction-by-simulation need not assume that predictions depend on time or resources, or that phonological prediction is less likely to occur compared to semantic or syntactic predictions.

Appendix A. (*continued*)

No.	Sentence (L1 cloze; L2 cloze)	Object names
	when the sun went behind the cloud, but the weather got better later. (100; 92)	[piero]/ bear[kuma]/ globe [tikyuugi]
6	The woman forgot to affix a stamp when posting the letter, and she got it back yesterday. (92; 82)	letter[tegami]/ lettuce [retasu]/ handcuff [tejyou]/ cat[neko]
7	The man didn't know the time because he forgot to wear the watch that he usually wears. (100; 92)	watch[tokei]/ washing machine[sentakuki]/ bird [tori]/ stamp[kitte]
8	The expensive wine is made from a special kind of grape that is grown only in the South of France. (92; 100)	grape[budou]/ grave [haka]/ pig[buta]/ comb [kusi]
9	To make sushi, the chef went to the market to buy some fish early in the morning. (100; 83)	fish[sakana]/ finger[yubi]/ dice[saikoro]/ elephant [zou]
10	To protect against an enemy's bullet or arrows, soldiers used to carry a shield all the time. (92; 67)	shield[tate]/ sheep[hituji]/ bamboo[take]/ giraffe [kurin]
11	The man was gathering honey, when he was stung by a bee and gave a cry. (100; 83)	bee[hati]/ bean[mame]/ flag[hata]/ tiger[tora]
12	The child believed that Santa Claus would come into her house down the chimney at midnight. (100; 83)	chimney[entotu]/ chick [hiyoko]/ pencil[enpitu]/ spoon[supu-n]
13	People can easily go to the island on foot since the government built a bridge last year. (100; 83)	bridge[hasi]/ brick[renga]/ ladder[hasigo]/ key[kagi]
14	The traveller went to the desert because he wanted to ride a camel and go exploring. (100; 100)	camel[rakuda]/ camera [kamera]/ racket[raketto]/ toothbrush[haburasi]
15	The woman found the room was too hot and humid, so to get some fresh air, she opened the window completely. (94; 100)	window[mado]/ windmill [huusya]/ match[matti]/ corn[toumorokosi]
16	The bird cannot fly because it injured its wing when it had a fight with another bird. (100; 83)	wing[hane]/ witch [majyo]/ nose[hana]/ candle[rousoku]

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(38), 419–439. http://dx.doi.org/10.1006/jmla.1997.2558.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264. http://dx.doi.org/10.1016/S0010-0277(99)00059-1.

Baayen, H. R., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. http://dx.doi.org/10.1016/j.jml.2007.12.005.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*(4), 457–474. http://dx.doi.org/10.1016/j.jml.2007.09.002.

Bates, D. M., Maechler, M., & Dai, B. (2008). lme4: Linear mixed-effects models using S4 classes.

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes, 22*(5), 633–660. http://dx.doi.org/10.1080/01690960601000746.

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology, 25*(5), 547–567. http://dx.doi.org/10.1080/20445911.2013.812093.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417–436. http://dx.doi.org/10.1016/j.jecp.2012.01.005.

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 35*(4), 1029–1040. http://dx.doi.org/10.1037/a0015901.

Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences, 10*(12), 564–570. http://dx.doi.org/10.1016/j.tics.2006.10.002.

De Groot, F., Huettig, F., & Olivers, C. N. L. (2016). When meaning matters: The temporal dynamics of semantic influences on visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 42*(2), 180–196. http://dx.doi.org/10.1037/xhp0000102.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*(1), 90–123. http://dx.doi.org/10.1016/0749-596X(91)90012-9.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121. http://dx.doi.org/10.1038/nn1504.

Dijkstra, T., Van Jaarsveld, H., & Brinke, S. Ten. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, 1, pp. 51–66, http://dx.doi.org/10.1017/S1366728998000121.

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*(5), 1461–1469. http://dx.doi.org/10.1037/a0036756.

Foucart, A., Ruiz-Tada, E., & Costa, A. (2016). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition, 19*(1), 213–219. http://dx.doi.org/10.1017/S1366728915000486.

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*(5), 862–877. http://dx.doi.org/10.1037/0278-7393.31.5.862.

Hintz, F., & Huettig, F. (2015). The complexity of the visual environment modulates language-mediated eye gaze. In R. Mishra, N. Srinivasan, & F. Huettig (Eds.), *Attention and vision in language processing* (pp. 39–55). Berlin: Springer. doi: 10.1007/978-81-322-2443-3_3.

Hintz, F., & Meyer, A. S. (2015). Prediction and production of simple mathematical equations: Evidence from visual world eye-tracking. *PLoS ONE, 10*(7), 1–17. http://dx.doi.org/10.1371/journal.pone.0130766.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, 1626*, 118–135. http://dx.doi.org/10.1016/j.brainres.2015.02.014.

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition, and Neuroscience, 31*(1), 19–31. http://dx.doi.org/10.1080/23273798.2015.1072223.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57*(4), 460–482. http://dx.doi.org/10.1016/j.jml.2007.02.001.

Ito, A., Corley, M., & Pickering, M. J. (2017). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 1–14. http://dx.doi.org/10.1017/S1366728917000050.

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language, 86*, 157–171. http://dx.doi.org/10.1016/j.jml.2015.10.007.

Ito, A., Martin, A. E., & Nieuwland, M. S. (2016a). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition and Neuroscience.* http://dx.doi.org/10.1080/23273798.2016.1242761.

Ito, A., Martin, A. E., & Nieuwland, M. S. (2016b). On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning Memory and Cognition.* http://dx.doi.org/10.1037/xlm0000315.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127*(2), 277–288. http://dx.doi.org/10.1016/j.actpsy.2007.06.003.

Kim, A. E., & Lai, V. T. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of Cognitive Neuroscience, 24*(5), 1104–1112. http://dx.doi.org/10.1162/jocn_a_00148.

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language, 61*(3), 326–338. http://dx.doi.org/10.1016/j.jml.2009.06.004.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning Memory and Cognition, 35*(2), 381–390. http://dx.doi.org/10.1037/a0014875.

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*(4), 574–588. http://dx.doi.org/10.1016/j.jml.2013.08.001.

McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology, 60*(5), 661–671. http://dx.doi.org/10.1080/17470210601183890.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475–494. http://dx.doi.org/10.1016/j.jml.2007.11.006.

Mishra, R. K., & Singh, N. (2014). Language non-selective activation of orthography during spoken word processing in Hindi-English sequential bilinguals: An eye tracking visual world study. *Reading and Writing, 27*(1), 129–151. http://dx.doi.org/10.1007/s11145-013-9436-5.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105–110. http://dx.doi.org/10.1016/j.tics.2006.12.002.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*, 329–392. http://dx.doi.org/10.1017/S0140525X12001495.

R Development Core Team (2015). A language and environment for statistical computing. R Foundation for Statistical Computing.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology. Human Perception and Performance, 37*(2), 514–528. http://dx.doi.org/10.1037/a0020990.

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia, 51*(3), 437–447. http://dx.doi.org/10.1016/j.neuropsychologia.2012.12.002.

Saslow, M. G. G. (1967). Latency for Saccadic Eye Movement. *Journal of the Optical Society of America, 57*(8), 1030. http://dx.doi.org/10.1364/JOSA.57.001030.

Segalowitz, N., & Hulstijn, J. H. (2009). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). New York: Oxford University Press.

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition, 128*(3), 302–319. http://dx.doi.org/10.1016/j.cognition.2013.02.013.

Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from eitherEllipsisor. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32*(2), 425–436. http://dx.doi.org/10.1037/0278-7393.32.2.425.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*(6), 557–580. http://dx.doi.org/10.1023/A:1026464108329.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences of the United States of America, 104*(30), 12530–12535. http://dx.doi.org/10.1073/pnas.0609927104.

Wheeldon, L. R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *The Quarterly Journal of Experimental Psychology, 44*(4), 723–761. http://dx.doi.org/10.1080/14640749208401307.

Wu, Y. J., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition, 129*(2), 418–425. http://dx.doi.org/10.1016/j.cognition.2013.08.005.