Automaticity and prediction in non-native language comprehension

Chapter · January 2021

DOI: 10.1075/bpa.12.02ito

2 authors:

Aine Ito
National University of Singapore

Martin J Pickering
The University of Edinburgh

Automaticity and prediction in non-native language comprehension

Aine Ito (Humboldt-Universität zu Berlin) & Martin J. Pickering

(University of Edinburgh)

A book chapter in “Prediction in Second Language Processing and Learning” (2021). Kaan, E. & Grüter, T. (Eds). John Benjamins Publishing

Abstract

Some evidence suggests that prediction is more limited in non-native language (L2) than native language (L1) comprehension. We evaluate the hypothesis that prediction is limited in L2 because prediction is largely non-automatic. We examine whether the subprocesses involved in prediction are *unconscious*, *unintentional*, *efficient* and *uncontrollable* (Bargh, 1994) to understand the extent to which prediction is automatic in L1 and L2. To unpack the subprocesses in prediction, we draw on Pickering and Garrod's (2013) proposal that people primarily use their production system for prediction, as well as a more automatic association-based mechanism. We conclude that at least some of the subprocesses in prediction are not fully automatic and suggest that these non-automatic processes can interfere with prediction in L2.

Introduction

People regularly predict upcoming information during language comprehension (Pickering & Gambi, 2018). By prediction in language comprehension, we mean pre-activation of information (e.g., semantic, syntactic, and phonological components of lexical representations) of an input (e.g., word, phrase) before the comprehender hears or reads the input. For example, upon hearing the verb *eat*, people may pre-activate the semantic representation [+edible] about an upcoming object. When a specific word (e.g., *cake*) is highly predictable, they may pre-activate the phonological (e.g., /keɪk/) or syntactic (e.g., a noun) components of that word's lexical representation. Such predictions likely facilitate efficient sentence processing by allowing people to access and process the information that has not yet occurred in the sentence. However, L2 speakers often predict more slowly or predict less detailed information than L1 speakers, as we review below. This lesser degree of prediction during comprehension might partly explain why L2 speakers have difficulty with online sentence processing. Thus, it is important to understand the mechanisms of prediction in L2 comprehension and how they might differ from those in L1 comprehension.

One possible explanation for the limitation in L2 prediction is that it is largely non-automatic. More generally, prediction is thought to involve both automatic processes (which are rapid and not subject to resource limitation) and non-automatic processes (which require time and resources) (Huettig, 2015; Pickering & Garrod, 2013). Pickering and Garrod (2013; see also Pickering & Gambi, 2018) proposed that the main mechanism people use for prediction is the mechanism they use to produce language. Since language production involves non-automaticity at each critical stage (conceptualization, lexical selection and grammatical encoding), prediction using the production system is also likely to be largely non-automatic. If so, L2 speakers may not have enough cognitive resources available for prediction, because L2 processing is generally more resource-demanding than L1 processing (Segalowitz & Hulstijn, 2009). We evaluate this hypothesis by examining the extent to which prediction during L2 processing is automatic.

This chapter falls into four sections. In the first section, we consider a graded view of automaticity (i.e., rather than all or none) in relation to "the four horsemen of automaticity" proposed by Bargh (1994). In the second section, we apply this notion to a model of language prediction. In the third and fourth sections, we review studies on prediction in L1 and L2 respectively, and reconcile their findings with the production-based prediction model (Pickering & Gambi, 2018; Pickering & Garrod, 2013).

Graded view of automaticity

Traditionally, automaticity has been seen as dichotomous: a process is either fully automatic or fully non-automatic (Shiffrin & Schneider, 1977). Automatic processes were considered to be unconscious, unintentional, efficient, and uncontrollable. Non-automatic processes were simply the opposite: conscious, intentional, inefficient, and controllable. However, this notion has been challenged. Some researchers (Bargh, 1994; Moors, 2016) have argued that processes are more or less automatic rather than fully automatic or non-automatic, because the features of automaticity are not necessarily co-present. That is, a process can be automatic in some respects but not in others. For example, a process can be uncontrollable but conscious. Thus, automaticity is seen as a graded notion according to this more recent view.

Bargh (1994) decomposed automaticity into four features, which he called “four horsemen”: *awareness*, *intentionality*, *efficiency*, and *controllability* (see Garrod & Pickering, 2007, for an application of this account to language production). *Awareness* refers to whether people are aware of the stimulus itself or its influence. *Intentionality* refers to whether people can instigate the process intentionally. *Efficiency* refers to whether the process draws on attentional resources. *Controllability* refers to whether people can stop or moderate the process once it has started. A process is regarded as most automatic if it occurs unconsciously, demands no attentional resources, and cannot be intentionally started or stopped. For example, masked priming is largely automatic in these respects, as participants are unaware of the stimuli (primes) and the priming effect occurs regardless of participants’ intention. As we review below, much of the evidence suggesting that prediction is non-automatic relates to efficiency, with some evidence suggesting that it is also non-automatic in terms of awareness, intentionality, and controllability.

A production-based model of prediction

In this section, we apply the four horsemen of automaticity proposed by Bargh (1994) to the mechanisms of prediction in the production-based prediction model (Pickering & Gambi, 2018; Pickering & Garrod, 2007, 2013). This model proposes that the central mechanism of prediction is *prediction-by-production*: people predict upcoming language with the mechanism they use to produce language. To do so, they covertly imitate the language they have comprehended, and derive the intention of the speaker while considering both linguistic and non-linguistic information (e.g., shared background knowledge). They may additionally apply adjustments to compensate for differences between the speaker and themselves, which allows them to predict what the speaker, rather than themselves, is likely to say. They then run the speaker’s intention through their own production system, retrieving the representations of what constitutes the upcoming utterance of the speaker. Prediction-by-production is relatively slow and requires resources because critical stages of language production require time and resources. People can also control and are aware of at least some of the production processes (e.g., lexical selection). Thus, prediction-by-production appears to involve processes that are non-automatic in terms of efficiency, awareness, intentionality, and controllability.

There is a complication to this account. Pickering and Garrod (2013) proposed that there is another aspect of prediction-by-production that is more automatic. This aspect of prediction involves a so-called forward model (see Wolpert, 1997). When speakers plan an utterance (e.g., *I will eat a cake*.), they also rapidly predict aspects of what they are going to say (e.g., the first phoneme /k/) using a forward model. This model involves a (learned) mapping between the speaker's intention and the predicted outcome. In the past, whenever the speaker has decided to say *cake*, she has subsequently uttered /*k*/ and experienced the sensation of uttering it, so the decision to utter *cake* now activates this sensation automatically. This prediction is used in self-monitoring: if the speaker now prepares the wrong phoneme (e.g., /*j*/), she notices that it mismatches the forward model and can revise her utterance.

When the comprehender predicts an utterance using prediction-by-production, he can also use the forward model, but this time using his covert imitation of the speaker's utterance (plus adjustments for differences between comprehender and speaker). After hearing *I will eat a*), he constructs the intention to produce the word *cake*, and at the same time predicts (say) /*k*/ via a forward model. Prediction using the forward model is highly automatized.

But at the same time, he begins to implement the process of produce the word *cake* and thereby generates predictions of what the speaker will say (Pickering & Gambi, 2018) – predicting its meaning, grammar, and sound in turn. This approach to prediction is slower and resource-intensive (as it is implemented in the language production system). For the purposes of this chapter, we focus on this latter aspect of prediction-by-production – in which the comprehender computes representations corresponding to the stages of language production (see Pickering & Gambi, 2018).

Prediction-by-production is complemented by another mechanism called *prediction-by-association*, which is not based on production. In this mechanism, lexical priming plays a role, and representations that are associated with words present in the context receive activation. It appears to be automatic in terms of efficiency and awareness, because lexical priming occurs very rapidly (Ratcliff & McKoon, 1981), and people do not seem to be aware of the process. It appears largely automatic in terms of controllability and intentionality as well (e.g., because people cannot stop it). There is some evidence that associations-based priming effects might be affected by strategic processing (Lau et al., 2013) but not at early stages (Grossi, 2006), suggesting that rapid spreading of activation is uncontrollable. While prediction-by-association is more rapid and consumes fewer resources, it is not as accurate as prediction-by-production, because words in the context often have lexical associations with words that are unlikely to follow the context. For example, the context *I will ride*… is likely to activate words representing a vehicle, but words like *motorbike* are not likely to occur if the speaker is a child. People take such information into account when predicting upcoming language (Kamide et al., 2003), but prediction-by-association activates lexically associated words even if they do not fit with the context. Thus, Kukona et al. (2011) found that when people heard *Toby will arrest*…, they tended to look at a correct theme (e.g., crook) as well as a typical agent of the verb (e.g., policeman) even though it was not a plausible object of the verb (see the next section for details of this study).

Automaticity of prediction in L1

We propose that prediction-by-production is largely non-automatic, as critical stages in language production involve non-automatic processes (Hartsuiker & Moors, 2017; Roelofs & Piai, 2011) – though note the discussion of forward modelling above. Language production involves several steps from conceptualization of the meaning that a speaker wants to convey, selection of a lexical item, morphological encoding, and retrieval of the phonological form up to articulation. Conceptualization is cognitively demanding (Bock, 1982), people are aware of the process, and they can decide to start or interrupt the process; hence it is largely non-automatic in terms of all four horsemen (Garrod & Pickering, 2007). Lexical selection seems non-automatic in terms of awareness and intentionality, as speakers are generally aware of their lexical choices and they can voluntarily instigate the process (Garrod & Pickering, 2007).

Processes of linguistic encoding such as syntactic planning, phonological word-form selection, and phoneme selection also seem to require working memory (e.g., Cook & Meyer, 2008; Ferreira & Pashler, 2002; Hartsuiker & Barkhuysen, 2006), and hence are non-automatic in terms of the horseman of efficiency. For example, Hartsuiker and Barkhuysen (2006) found more number agreement errors (e.g., *The coupon in the flyers were*…) when participants had to remember a word list than when they did not, suggesting that syntactic planning in speech production requires working memory resources and is therefore not fully automatic (see also Fayol et al., 1994). Experiments using a dual-task (picture naming and tone discrimination) found that picture naming latency was shorter when it followed a high versus low cloze context, when the lexical frequency was high versus low (Ferreira & Pashler, 2002), and when a distractor presented with the target picture was phonologically related to the target than when it was phonologically unrelated (Cook & Meyer, 2008). Critically, these effects carried forward to the following tone discrimination task, suggesting that lemma selection (affected by cloze), word-form selection (affected by frequency), and phoneme selection (affected by phonological relatedness) all require cognitive resources (because, if not, they should not affect the tone-discrimination task) and are non-automatic in terms of efficiency. Some subprocesses are treated as automatic in some production models, such as spreading activation between lemma and lexeme (Dell, 1986) or formulating and articulatory procedures (Levelt, 1989). But critically, the above findings suggest that each key stage of production involves some non-automaticity.

Similarly, there is evidence that prediction is not fully automatic in terms of efficiency, and whether prediction occurs is subject to time and resources available during comprehension. For example, Huettig and Janse (2016) investigated the effect of people's working memory span and general processing speed on their predictive eye movements using the visual world paradigm, where people listen to language while viewing some objects. In their study, L1 Dutch speakers listened to Dutch sentences such as *Kijk naar de afgebeelde piano* ('Look at the$_{\text{-common}}$ displayed piano$_{\text{-common}}$.') while viewing a scene containing four objects. Articles in Dutch are gender-marked (e.g., *de* is common gender), and the gender of only the target object name matched the article's gender. As expected, participants fixated the target object predictively (i.e., before the target noun was mentioned). Critically, their working memory and processing speed (measured using multiple verbal and non-verbal tasks) were positively related to the degree of predictive eye movements: Participants who had a greater working memory span and a higher processing speed made more predictive eye movements. These effects suggest that some processes involved in making predictive eye movements require resources, and hence are non-automatic with respect to efficiency in the four horsemen.

The findings in Huettig and Janse (2016) are strengthened by evidence showing a causal link between working memory and predictive eye movements. (Ito, Corley, et al., 2018; Experiment 1) found that an additional working memory load interfered with predictive eye movements. L1 English speakers listened to predictable sentences (i.e., with a constraining verb) such as *The lady will fold the scarf* or unpredictable sentences (i.e., with a less constraining verb) such as *The lady will find the scarf* while viewing a scene with four objects. Only the target object (here, scarf) was plausible after the verb in predictable sentences (and hence it was predictable given that one of the pictures would be described), whereas all four objects were plausible after the verb in unpredictable sentences. The participants' task was to click on the mentioned object. Half of the participants performed an additional working memory task to tax their working memory (i.e., to make less cognitive resources available) during sentence processing. They saw five abstract words before the listen-and-click task and recalled the words immediately after the task. Participants fixated the target object predictively in predictable sentences, but these predictive eye movements were delayed in those who performed the working memory task compared to those who did not. This finding, together with the findings in Huettig and Janse (2016), suggests that cognitive resources are required for making predictive eye movements.

There is also evidence that prediction is dependent on time available for generating predictions, which also suggests that prediction is non-automatic in terms of efficiency. Huettig and Guerra (2019) used the scenes and sentences from Huettig and Janse (2016) and investigated effects of speech rate and preview time on predictive eye movements. When participants viewed the objects for four seconds before the sentence began, they fixated the target predictively both when the speech was at a normal rate (the sentence took just under 2 seconds) and at a slow rate (it took just over 4 seconds). However, when they were given a one-second preview, predictive eye movements occurred only in the slow rate condition. With the shorter preview, it could be that participants were still in the process of encoding of the scene, which then interfered with generating predictions.

Similarly, Ito et al. (2016) found that people do not pre-activate full representations of a predictable word when they do not have enough time to generate predictions. L1 English speakers read predictive sentence contexts (e.g., *The student is going to the library to borrow a*…), followed by a predictable word (*book*), a word that was form-related (*hook*) or semantically related (*page*) to the predictable word, or an unrelated word (*sofa*). Participants read the sentences word-by-word at a 500 ms or 700 ms presentation rate. Consistent with previous studies (Federmeier & Kutas, 1999; Laszlo & Federmeier, 2009), Ito et al. found that all unexpected words elicited an N400 effect relative to the predictable word, but words that were form-related or semantically related to the predictable word elicited a reduced N400 effect relative to the unrelated word. The N400 reduction was greater when the word predictability was higher (estimated in a cloze test, where people completed sentence fragments with a continuation that came to mind), suggesting that people pre-activated word form and meaning of the predictable word. Critically, Ito et al. found the N400 reduction for semantically related words at both presentation rates, but the N400 reduction for form-related words emerged only at the slower presentation rate (700 ms). This finding suggests that prediction of the phonological component of lexical representations takes time; when time is limited, the semantic component may be predicted but not the phonological component.

The effects of working memory and time on prediction suggest that some processes involved in prediction are not automatic in terms of efficiency. But which predictive processes require time and cognitive resources? There is evidence indicating that adjusting the perspective to the speaker in prediction may be cognitively demanding. Corps et al. (2021) tested prediction of an upcoming object in sentences like *I would like to wear the nice … tie/dress*, where the continuation is likely to be different depending on the gender of the speaker (e.g., a male speaker might say *tie*, whereas a female speaker might say *dress*). That is, the listener must adjust for differences in perspective to predict what the speaker is likely to say appropriately. L1 English speakers heard these sentences in a male or female voice (the continuation always matched the speaker's gender) while viewing objects that did or did not match the speaker's gender (e.g., a tie, a dress) together with two distractors (e.g., a drill, a hairdryer). Upon hearing the verb, they fixated objects that were compatible with the verb (e.g., a tie, a dress) over the distractors irrespective of the speaker's gender. They also predictively fixated objects that were expected based on the speaker's gender (e.g., a tie for a male speaker), but these eye movements occurred later than association-based predictive eye movements. The delayed effect for speaker-specific prediction suggests that adjustments for speaker-listener perspective differences require resources.

So far, we have discussed automaticity in terms of efficiency. What about the other horsemen? Brothers et al. (2017) investigated to what extent prediction generation could be modulated by people's strategies (hence tapping into intentionality in the four horsemen). In their ERP study (Experiment 1), participants read medium-cloze (cloze = 50.7%) or low-cloze (0.7%) sentences, and they were instructed to simply comprehend the sentences or to try to predict the final word of each sentence. They found an N400 effect followed by a frontal positivity for low-cloze words relative to medium-cloze words, replicating previous findings (Federmeier & Kutas, 1999; Van Petten & Luka, 2012). Critically, these effects were larger when participants were instructed to predict than not, suggesting that participants' strategies can modulate effects of sentence constraints.

Brothers et al. (2017) additionally tested whether reliability of probabilistic cues affect prediction in a self-paced reading task (Experiment 2). The proportion of trials where a highly predictive sentence context was followed by the predictable word was varied across participants. Participants read target words faster when they followed a predictive context versus a non-predictive context. This predictability effect depended on the reliability of the cues; participants who read most of the predictive contexts followed by an unexpected word did not show this effect, and this effect was largest in participants who read most of the predictive contexts followed by the expected word (see also Hopp, 2016, for consistent findings in a visual world study). This suggests that people can inhibit prediction. The findings in Brothers et al. (2017) together demonstrate that people's strategies affect the degree of prediction, and prediction is not automatic in terms of intentionality or controllability.

In contrast to prediction-by-production, prediction-by-association seems largely automatic in terms of efficiency in L1 processing, because it is driven by lexical priming, and spreading of activation occurs rapidly in lexical priming (Neely, 1991; though later stages might be subject to strategic processing, Lau et al., 2013). Findings suggesting that prediction-by-association occurs rapidly comes from Kukona et al. (2011). In their Experiment 1, L1 English speakers listened to sentences with a semantically constraining verb such as *Toby arrests the crook*, while viewing images that corresponded to a typical patient of the constraining verb (the target: crook), a typical agent of the verb (the competitor: policeman), the mentioned subject (Toby), and unrelated distractors (surfer, florist). After hearing the verb (*arrests*), participants fixated not only the predictable target but also the unexpected competitor more than unrelated distractors. Participants started to fixate the target more than the competitor only after hearing the target. This finding suggests that predictive eye movements are (at least partly) driven by thematic priming (Ferretti et al., 2001) – priming from a verb to its typical agent and patient (irrespective of whether the agent/patient plausibly follows the verb). In Experiment 2, they used passive sentences such as *Toby was arrested by the policeman* and similarly found that participants fixated both the target (e.g., a policeman) and the competitor (e.g., a crook) over distractors after hearing the verb. But unlike in Experiment 1, they found that participants fixated the target more than the competitor before hearing the target. This finding suggests that the predictive eye movements are also driven by contextual information (i.e., not only by thematic priming), but context-based prediction may take more time to be generated. Crucially, thematic priming contributed to the early stage of predictive eye movements, suggesting that association-based prediction occurs rapidly.

Note that Gambi et al. (2016) used similar active sentences as those in Kukona et al.'s (2011) Experiment 1 and found no fixation bias towards an unexpected typical agent of a verb. After hearing an agent followed by a verb (e.g., *Pingu will ride*), L1 adults and children looked at appropriate patients (e.g., a horse) but not an inappropriate agent (e.g., a cowboy). As the authors discuss, one of the differences between the two studies was that the speech rate was much slower in Gambi et al. than in Kukona et al. Thus, it could be that participants in Gambi et al. had enough time to use the contextual information for prediction and predicted an appropriate referent by considering the syntactic structure (i.e., the upcoming referent is going to be a patient of the verb).

Prediction-by-association appears largely automatic in terms of controllability because people cannot stop the process. However, there is also evidence that people's strategies affect the degree of semantic priming (Lau et al., 2013), suggesting that spreading of activation may not be fully automatic in terms of intentionality or controllability (at least at late stages, see the discussion above). Kukona et al. (2016) found that people with higher comprehension skills (measured in several different tests) were better at inhibiting implausible referents that shared features with predictable referents (e.g., showing fewer looks to a white car after hearing *The boy will eat the white*…). People with higher comprehension skills may be better at controlling activation of inappropriate information (see also Peters et al., 2018), suggesting that prediction-by-association may not be fully automatic in terms of controllability or intentionality.

Automaticity of prediction in L2

Language comprehension is generally more resource-demanding (i.e., less automatic in terms of efficiency) in L2 than in L1 (Segalowitz & Hulstijn, 2009). L2 speakers may be slower or less likely to predict than L1 speakers because it takes more time and resources for L2 speakers to access lexical representations and build up sentence meaning that they can use for prediction. Language production in L2 also appears to be non-automatic in terms of efficiency because L2 production (picture naming) requires more time than L1 production (Ivanova & Costa, 2008), and lexical selection and monitoring processing are affected by task demands in L2 production (Declerck & Kormos, 2012). Consequently, prediction-by-production is likely to be less automatic in L2 than in L1, and L2 speakers may not predict as efficiently as L1 speakers. In addition, prediction-by-association in L2 may be less efficient than that in L1, because spreading of activation may be slower and lexical links may be weaker in L2 than in L1. In line with this, Dijkgraaf et al. (2019) found that predictive eye movements based on semantic associations were slower and weaker in L2 than in L1. In sum, non-automaticity involved in comprehension and production are both likely to account for reduced prediction in L2 speakers.

We discuss why prediction in L2 is limited compared to prediction in L1 in relation to (non-)automaticity in L2 processing. The first possibility is that similarity between L1 and L2 modulates prediction in L2 (Dussias et al., 2013; Foucart et al., 2014). Processing of lexical or grammatical features that are absent in L1 may be particularly subject to resource limitations because rules or representations in L1 (which L2 speakers can arguably apply more automatically) cannot be applied. Thus, L2 speakers may have difficulty in using L2-specific linguistic features for prediction (Bobb et al., 2015; Foucart & Frenck-Mestre, 2011). Another possibility is that L2 speakers may be slower to pre-activate L2 words because of interference from lexical representations in their L1. Lexical representations in L1 may compete when L2 speakers access L2 words in comprehension (Spivey & Marian, 1999) and production (Costa et al., 2006; Costa & Santesteban, 2004). For example, L2 speakers may try to pre-activate phonological representations in both L1 and L2 even when comprehending sentences in L2, and this may delay access to phonological representations in L2 and make prediction of phonology less efficient. However, there are other factors that may account for L1-L2 differences in the degree of prediction, some of which are unrelated to automaticity such as differences in frequency biases. For example, L1 and L2 speakers are likely to differ in how often they are exposed to particular combination of words, so they may predict different words (see Kaan, 2014). In this review, we discuss accounts that are related to automaticity (cf. Hopp, 2010, 2018).

Several studies have shown L2 speakers' difficulty in making predictions by using linguistic features that do not occur in their L1. Lew-Williams and Fernald (2010) investigated whether L2 Spanish speakers (with L1 English) use grammatical gender cues efficiently to identify target objects like L1 speakers in the visual world paradigm. Participants listened to sentences such as *Encuentra la pelota*. ('Find the$_{\text{-fem}}$ ball$_{\text{-fem}}$.'), while viewing two objects that had either the same or different gender in Spanish. L1 speakers fixated the target object more quickly when two objects had a different gender (i.e., when the article gender matched only the target object) compared to when they had the same gender, suggesting that L1 speakers used the grammatical gender cues efficiently to identify the target object. In contrast, L2 speakers did not benefit from the gender cues; they were no quicker to fixate the target object when the article gender allowed them to narrow down possible referents to a single object than when it did not. Thus, L2 speakers may not use gender (which is absent in their L1) to facilitate processing of upcoming words to the same extent as L1 speakers do.

Using German, Hopp (2013) found that L2 German speakers (with L1 English) predicted upcoming objects based on a gender cue, but only L2 participants who produced gender correctly 97% of the time in another production task predicted as quickly as L1 speakers. These participants were near-native German speakers who had been exposed to German for 42 years on average. L2 participants who produced gender correctly 77% of the time showed delayed predictive eye movements, and these participants had been exposed to German for 24 years on average. Hence, extremely high L2 proficiency may be needed for L2 speakers to predict like L1 speakers (as suggested by the performance in the production task). In a follow-up study, Hopp (2016) showed that L2 speakers came to use gender for prediction after being trained on gender assignment. The delayed prediction in less proficient L2 speakers in Hopp (2013) or the lack of prediction in L2 speakers in Lew-Williams and Fernald (2010) may be explained by L2 speakers' difficulty in encoding grammatical gender (as the production task results suggest), or their difficulty in using gender efficiently, or both.

Mitsugi and MacWhinney (2016) investigated prediction using case markers in Japanese (cf. Kamide et al., 2003; Experiment 3) in L1 Japanese speakers and L2 Japanese speakers (with L1 English) who had good knowledge of Japanese case markers in an offline test (with 86.5% accuracy). Participants listened to Japanese sentences with either a dative structure (e.g., *Gakkou-de majimena*$_{\text{-Adj}}$ *gakusei-ga*$_{\text{-Nom}}$ *kibishii*$_{\text{-Adj}}$ *sensei-ni*$_{\text{-Dat}}$ *shizukani*$_{\text{-Adv}}$ *tesuto-o*$_{\text{-Acc}}$ *watashita*$_{\text{-V}}$.; 'At the school, the serious student quietly handed over the exam to the strict teacher.') or an accusative structure (e.g., *Gakkou-de majimena*$_{\text{-Adj}}$ *gakusei-ga*$_{\text{-Nom}}$ *kibishii*$_{\text{-Adj}}$ *sensi-o*$_{\text{-Acc}}$ *shizukani*$_{\text{-Adv}}$ *karakatta*$_{\text{-V}}$.; 'At the school, the serious student quietly teased the strict teacher.'). A dative NP is almost always followed by an accusative NP, so the occurrence of a theme is predictable (Kamide et al., 2003). L1 speakers showed increased looks to a theme (e.g., a test) in the dative structure relative to the accusative structure before the theme was mentioned, suggesting that they predicted the occurrence of a theme. In contrast, L2 speakers' eye movements were not affected by the case marker, and they directed their looks to the theme only after it was mentioned. These results also suggest that L2 speakers may not use linguistic cues that are absent in their L1 for prediction as efficiently as L1 speakers.

The more direct evidence for the role of L1-L2 similarity comes from studies showing that L2 speakers whose L1 has grammatical gender are more likely to use gender cues for prediction than L2 speakers whose L1 does not have grammatical gender. Dussias et al. (2013) tested the use of gender cues in L1 Spanish speakers and L2 Spanish speakers with L1 English or L1 Italian. Replicating Lew-Williams and Fernald (2010), they found that L1 speakers used gender cues to quickly identify the upcoming noun. L1 Italian speakers used the feminine article but not the masculine article (for which there seems no clear explanation). L1 English speakers with a similar L2 proficiency did not use articles, but those who with a higher proficiency (measured in grammar and vocabulary tests) showed predictive eye movements. Their findings suggest that both L1-L2 similarity and L2 proficiency modulate prediction.

Foucart et al. (2014) investigated prediction during reading in L1 Spanish speakers and L2 Spanish speakers (with L1 French) using Spanish highly predictable sentence contexts (e.g., *El pirata tenía el mapa secreto, pero nunca encontró* …; 'The pirate had the secret map, but he never found…') which were followed by either the expected continuation (e.g., *el tesoro*-masc; 'the treasure') or an unexpected but plausible continuation (e.g., *la gruta*-fem; 'the cave'). If participants predict the expected noun, the unexpected article was incompatible with the noun, so it was expected to elicit a different ERP than the expected article (cf. Wicha et al., 2004). In this study, both L1 and L2 speakers showed more negative ERPs for unexpected articles than for expected articles. This effect did not differ across groups, suggesting that L2 speakers predicted to a similar extent as L1 speakers. The negativity for unexpected versus expected articles was also found during listening comprehension in both L1 Spanish speakers (Foucart et al., 2015) and L2 Spanish speakers (with L1 French) (Foucart et al., 2016). Like Dussias et al. (2013), Foucart et al. (2016) argued that language similarity between L1 and L2 may mediate prediction (see also Foucart, this volume, for more detailed discussion).

Ito et al. (2021) examined prediction based on verb constraints that are specific to the L2 (Vietnamese) in L1 Vietnamese speakers and L2 Vietnamese speakers (with L1 German). For example, the Vietnamese verb *bê* means to carry a heavy object with two hands, while its German translation equivalent *tragen* means to carry (not necessarily a heavy object or with two hands), so the verb has a different mapping between Vietnamese and German. If L2 speakers suffer from interference from German verb constraints (i.e., L1 transfer effects), they might predict anything that one can carry when they hear *bê*. Participants listened to

Vietnamese sentences (e.g., *Nam bê một chồng sách*; 'Nam carries a [classifier] book') which contained either a different-mapping (e.g., *bê – tragen*) or a similar-mapping verb (e.g., *ăn – essen*; both mean 'to eat') between Vietnamese and German. Upon hearing the verb, both L1 and L2 speakers fixated objects that met the verb constraints. But critically, L2 speakers also fixated objects that did not meet the Vietnamese verb constraints but met the verb constraints of the German translation equivalent (e.g., light things that can be carried). Thus, L2 speakers may have difficulty using L2 specific lexical features for prediction (see also Van Bergen & Flecken, 2017) and this may explain the lesser degree of prediction in L2 speakers.

It is important to consider whether L2 speakers' difficulty using L2-specific lexical or grammatical features for prediction comes from non-automaticity (e.g., inefficiency) of processing these features or incomplete acquisition/different representations of these features. The former possibility assumes that L2 speakers know lexical representations or grammatical rules in L2 but cannot use them in an automatic manner like L1 speakers, which could be because they often learned these features via explicit learning (e.g., by memorizing rules) (Hulstijn, 2002). As application of the grammatical rules are not as automatized as in L1 speakers, L2 speakers may not be able to use these features before predictable input comes up. The latter possibility assumes that the information that L2 speakers can use is different from what L1 speakers use. For example, if L2 speakers of Spanish do not have a lexical representation that the noun *ball* is feminine in Spanish, they cannot determine that it is grammatical after the feminine article *la*. As we reviewed above, there are cases where L2 speakers do not predict using the grammatical rules they know fairly well, as an offline test suggests (Mitsugi & MacWhinney, 2016). Nevertheless, L2 speakers seem more likely to predict using L2 specific lexical/grammatical features as their proficiency increases (Dussias et al., 2013; Hopp, 2013, though effects of proficiency may not be robust at all levels, cf. Hopp, 2015). Thus, automaticity in using L2 specific lexical/grammatical features online seems to be enhanced as L2 proficiency increases.

L2 speakers seem to predict less than L1 speakers even when L2 specific lexical/grammatical features are not primary cues for prediction. In an ERP study, DeLong et al. (2005) examined context-based prediction by presenting L1 English speakers with a predictive context (e.g., *The day was breezy so the boy went outside to fly* ...) followed by either the expected continuation (e.g., *a kite*) or an unexpected but plausible continuation (e.g., *an airplane*). The expected word began with a consonant and the unexpected word began with a vowel, or vice versa. If people predict the expected word (e.g., *kite*), encountering an unexpected article (e.g., *an*) was expected to interfere with processing. They found that the N400 amplitude at the article was smaller (more positive) when the cloze probability of the article was higher. Because this effect was found before the presentation of the noun, it was taken as evidence for prediction of the expected noun. It is important to note that the *a*/*an* article effect has not replicated in later studies (Ito et al., 2017b; Nieuwland et al., 2018). Notably, Nieuwland et al. (2018) conducted a large-scale replication (*N*=334) of DeLong et al. (2005) and did not find any effect at the article (but see Nicenboim et al., 2020 for a meta-analysis showing a small article effect).

Using this design, Martin et al. (2013) examined context-based prediction in L1 English speakers and L2 English speakers (with L1 Spanish). L1 speakers showed a negativity for the unexpected versus expected article 250 – 400 ms after the article onset. However, L2 speakers showed no effect at the article, though they were familiar with the *a*/*an* rule and showed an N400 effect for the unexpected versus expected nouns like L1 speakers. However, given the replication failures discussed above, it is unclear to what extent the L1-L2 difference in this study demonstrates evidence for limited prediction in L2 speakers.

Ito, Pickering, et al. (2018) tested context-based prediction in L1 English speakers and L2 English speakers (with L1 Japanese) using the visual world paradigm. Participants listened to sentences containing a highly predictable word (e.g., *The tourists expected rain when the sun* *went behind the* <u>*cloud*</u>*,* …) while viewing four objects. The critical object represented the predictable word (e.g., *cloud*), an English competitor which shared initial phonemes with the predictable word (e.g., *clown*), or an unrelated word (e.g., *globe*) (We will discuss another condition later in this section.). L1 speakers fixated both target and English competitor objects over the unrelated object before the target word was mentioned, suggesting that they predicted the phonological form of the target word. L2 speakers fixated the target predictively but they fixated the English competitor only well after the target word was mentioned. Thus, L2 speakers predicted some information about the target word, but there was no evidence that they predicted the phonological information. Since the target words were predictable from the context, L2 speakers seem to predict less than L1 speakers even when L2 specific lexical or grammatical features are not a primary cue for prediction.

Similarly, Ito et al. (2017a) investigated prediction of word form and meaning in L2 English speakers (with L1 Spanish) using the design in Ito et al. (2016; see the previous section for details) and found no evidence for prediction in L2 speakers. They showed an N400 reduction for words that were semantically related to the predictable word relative to unrelated words. However, this effect was no stronger when the cloze probability was higher, which is inconsistent with a prediction account. Form-related words showed a similar N400 to unrelated words, and thus there was no evidence that L2 speakers predicted word form. L2 speakers showed a Late Positive Complex (LPC) effect for form-related words relative to unrelated words at a slower word presentation rate (700 ms per word). This effect was larger in sentences with higher than lower cloze values, suggesting that L2 speakers processed sentences incrementally, combining both top-down information (sentence meaning) and bottom-up information (form-similarity). Nevertheless, they did not predict like L1 speakers (i.e., there was no N400 reduction for form-related words). Together, these findings (Ito et al., 2017a; Ito, Pickering, et al., 2018; Martin et al., 2013) indicate that L2 speakers may not predict like L1 speakers even when predictions are not based on L2 specific knowledge.

We mentioned that another factor that could affect prediction in L2 speakers is interference from L1 (Costa et al., 2006; Spivey & Marian, 1999). Given that lexical representations in L1 may interfere even when the L1 is irrelevant (e.g., during L2 comprehension), such interference may occur largely automatically (e.g., in terms of controllability) and may interfere with prediction in L2 by delaying access to L2 words (Hopp, 2018). In the study discussed above, Ito, Pickering, et al. (2018) tested whether L2 English speakers (with L1 Japanese) predict phonological information of a highly predictable word in their L1 by including a Japanese competitor condition, in which the critical object name shared initial phonemes with the predictable word when translated into Japanese (e.g., competitor: *bear*; Japanese: *kuma*; target: *cloud*; Japanese: *kumo*). If co-activating the Japanese equivalent of the target word interfered with prediction, we would have expected more fixations on the Japanese competitor than the unrelated object, but we did not find such an effect. However, this study did not find any evidence for word form prediction in L2 speakers, so it is unclear whether L2 speakers predict form in L1 when they do so in L2. If they do, the lesser degree of prediction in L2 may arise not only from resource limitations but also from interference from the L1. Regarding this issue, there is evidence suggesting that top-down language control affects lexical co-activation, such that lexical access to L1 words during L2 comprehension is facilitated by increased exposure to the L1 (Hoversten & Traxler, 2020). An interesting question for future research is whether lexical activation in L1 occurs automatically when people comprehend predictive sentences in L2 (cf. Ito et al., 2021) and whether that interferes with prediction in L2.

This review of studies suggests that whether L2 speakers predict also depends on the type of information involved. In fact, many studies found evidence for prediction of semantic information (Chambers & Cooke, 2009; Chun & Kaan, 2019; Dijkgraaf et al., 2017; Ito, Corley, et al., 2018). Some studies found evidence for prediction using syntactic information (Foucart et al., 2014, 2016; Hopp, 2013) while others did not (Lew-Williams & Fernald, 2010; Mitsugi & MacWhinney, 2016). In an investigation of semantic and syntactic prediction, Hopp (2015) found that L2 speakers used semantic information but not syntactic information for prediction. Upon hearing a verb in SVO sentences like *Der*$_{\text{-Nom}}$ *Wolf tötet*$_{\text{-V}}$ *gleich den*$_{\text{-Acc}}$ *Hirsch* ('The wolf will soon kill the deer'), L2 German speakers (with L1 English) looked at an appropriate patient (e.g., a deer) before it was mentioned, suggesting that they used verb semantics to predict an upcoming referent. However, they also tended to look at the same image (e.g., a deer) in OVS sentences like *Den*$_{\text{-Acc}}$ *Wolf tötet*$_{\text{-V}}$ *gleich den*$_{\text{-Nom}}$ *Jäger*. ('The hunter will soon kill the wolf'), suggesting that they did not use the case marker to predict that the upcoming referent would be an agent (rather than a patient) of the verb. Grüter et al. (2020) investigated prediction based on classifiers in Chinese and found that L2 speakers (with varied L1) tended to be distracted by objects that were grammatically inappropriate but matched a semantic class of a classifier (e.g., classifier *tiao* – long and flexible). This finding suggests that L2 speakers may be more likely to use semantic information (a semantic class of a classifier) than syntactic information (grammatical match between a classifier and a noun) for prediction. However, the studies that investigated prediction of phonological or orthographic word form have never found evidence for prediction in L2 speakers (Ito et al., 2017a; Ito, Pickering, et al., 2018; Martin et al., 2013). Thus, L2 speakers often use semantic information for prediction, while they are less likely to use syntactic information, and may not use word form information.

This pattern can be explained by prediction-by-production models (Pickering & Gambi, 2018; Pickering & Garrod, 2013), because in production, people first activate the semantic information of the word they are going to produce (conceptualization stage), then the syntactic information (grammatical encoding stage), and finally the phonological information (phonological encoding stage) (e.g., Levelt et al., 1999). Prediction-by-production involves the same stages, but the later stages may not occur when there is not enough time or resources available (e.g., prediction stages may be “cut-off” after the syntactic stage if time does not allow people to proceed to the phonological stage), because each stage in production entails some non-automaticity (e.g., Ferreira & Pashler, 2002; Garrod & Pickering, 2007; Hartsuiker & Barkhuysen, 2006). Thus, under this account, prediction of semantic information is more likely to occur than prediction of syntactic information, and prediction of syntactic information is more likely to occur than prediction of word form. Prediction-by-production during L2 processing may be particularly subject to resource limitations because of fewer resources available during L2 processing compared to L1 processing, making it less likely for L2 speakers to complete the full stages of prediction-by-production.

In the domain of L2 prediction research, there is not much evidence that relates to intentionality or controllability in the four horsemen (i.e., to what extent L2 speakers can control predictive processing). Regarding awareness, Curcic et al. (2019) used a Fijian-based miniature language and tested whether L2 speakers used a gender-marking rule to predict an upcoming noun based on the gender of a determiner. They found that L2 speakers who were aware that determiners could be used for prediction were more likely to predict than those who were unaware. This finding suggests a role of awareness, and prediction in L2 is non-automatic in terms of awareness.

Regarding efficiency, Ito, Corley, et al. (2018; Experiment 2) repeated the experiment investigating effects of working memory on predictive eye movements with L2 speakers of English (see the previous section for details) and found similar effects of working memory in L2 speakers. L2 speakers (with varied L1s) made predictive eye movements when there was no working memory task, but the concurrent working memory task delayed these eye movements. This finding suggests that working memory mediates predictive eye movements in L2 speakers and prediction is non-automatic in terms of efficiency – in line with the findings in L1 speakers. This finding is also consistent with Kaan (2014), who argued that the same factors (e.g., working memory) mediate prediction in L1 and L2.

Conclusion

We evaluated how automatic prediction is by reviewing studies on prediction in L1 and L2 speakers. A large body of evidence suggests that prediction requires time and resources, and it tends to be reduced in L2 speakers, who have fewer resources available during comprehension compared to L1 speakers. These findings suggest that prediction is non-automatic in terms of efficiency in the four horsemen. There is also evidence suggesting that people can control predictive processing, at least in L1, suggesting that prediction is non-automatic also in terms of intentionality and controllability. We suggest that prediction-by-association, which complements prediction-by-production, occurs rapidly and is largely automatic in terms of efficiency. However, evidence for such prediction is rather scarce, and further investigation will be required to separately assess how automatic prediction-by-association is in both L1 and L2 comprehension. Our review also suggests that L2 speakers often predict semantic information like L1 speakers in simple sentences and can sometimes predict syntactic information but may be slower to predict or may not predict word form information. The available evidence is in line with the production-based prediction accounts: Because production requires time and resources, the later stages of prediction-by-production (syntactic encoding and phonological encoding) may not occur when there are fewer cognitive resources available, as is the case in L2 comprehension.

References

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes; Applications* (pp. 1–40). Lawrence Erlbaum Associates, Inc.

Bobb, S. C., Kroll, J. F., & Jackson, C. N. (2015). Lexical constraints in second language learning: Evidence on grammatical gender in German. *Bilingualism: Language and Cognition*, *18*(3), 502–523. https://doi.org/10.1017/S1366728914000534

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*(1), 1–47. https://doi.org/10.1037/0033-295X.89.1.1

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, *93*, 203–216. https://doi.org/10.1016/j.jml.2016.10.002

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. https://doi.org/10.1037/a0015901

Chun, E., & Kaan, E. (2019). L2 Prediction during complex sentence processing. *Journal of Cultural Cognitive Science*, *3*(2), 203–216. https://doi.org/10.1007/s41809-019-00038-0

Cook, A. E., & Meyer, A. S. (2008). Capacity demands of phoneme selection in word production: New evidence from dual-task experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 886–899. https://doi.org/10.1037/0278-7393.34.4.886

Corps, R. E., Brooke, C., & Pickering, M. J. (2021). Do comprehenders predict from their own perspective? Manuscript in preparation

Costa, A., Heij, W. La, & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, *9*(02), 137. https://doi.org/10.1017/S1366728906002495

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, *69*(March), 42–71. https://doi.org/10.1111/lang.12321

Declerck, M., & Kormos, J. (2012). The effect of dual task demands and proficiency on second language speech production. *Bilingualism*, *15*(4), 782–796. https://doi.org/10.1017/S1366728911000629

Dell, G. S. (1986). A spreading-activation theory of retreival in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

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. https://doi.org/10.1038/nn1504

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, *20*(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, *34*(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469

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*(02), 353–387.

https://doi.org/10.1017/S0272263112000915

Fayol, M., Largy, P., & Lemaire, P. (1994). Cognitive overload and orthographic errors: When cognitive overload enhances subject–verb agreement errors. A study in French written language. *The Quarterly Journal of Experimental Psychology Section A*, *47*(2), 437–464. https://doi.org/10.1080/14640749408401119

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469–495. https://doi.org/10.1006/jmla.1999.2660

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(6), 1187–1199.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*(4), 516–547. https://doi.org/10.1006/jmla.2000.2728

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*(03), 379–399. https://doi.org/10.1017/S136672891000012X

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. https://doi.org/10.1037/a0036756

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). How do you know I was about to say “book”? Anticipation processes affect speech processing and lexical recognition. *Language, Cognition and Neuroscience*, *30*(6), 768–780. https://doi.org/10.1080/23273798.2015.1016047

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*(01), 213–219. https://doi.org/10.1017/S1366728915000486

Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers’ linguistic predictions. *Cognition*, *157*, 340–351. https://doi.org/10.1016/j.cognition.2016.10.003

Garrod, S., & Pickering, M. J. (2007). Automaticity of language production in monologue and dialogue. In A. S. Meyer, L. R. Wheeldon, & A. Krott (Eds.), *Automaticity and Control in Language Processing* (pp. 1–20). Psychology Press. https://doi.org/10.4324/9780203968512-7

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology*, *43*(1), 21–30. https://doi.org/10.1111/j.1469-8986.2006.00383.x

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: the role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, *35*(2), 221–234. https://doi.org/10.1080/23273798.2019.1648840

Hartsuiker, R. J., & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*(1–3), 181–204. https://doi.org/10.1080/01690960400002117

Hartsuiker, R. J., & Moors, A. (2017). On the automaticity of language processing. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge.* (pp. 201–225). American Psychological Association. https://doi.org/10.1037/15969-010

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*(4), 901–931. https://doi.org/10.1016/j.lingua.2009.06.004

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, *53*(3), 277–306. https://doi.org/10.1515/iral-2015-0014

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, *32*(2), 277–307. https://doi.org/10.1177/0267658315624960

Hopp, H. (2018). The Bilingual Mental Lexicon in L2 Sentence Processing. *Second Language*, *17*, 5–27. https://doi.org/10.11431/secondlanguage.17.0_5

Hoversten, L. J., & Traxler, M. J. (2020). Zooming in on zooming out: Partial selectivity and dynamic tuning of bilingual language control during reading. *Cognition*, *195*(October 2019), 104118. https://doi.org/10.1016/j.cognition.2019.104118

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

Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research, 1706*, 196–208. https://doi.org/10.1016/J.BRAINRES.2018.11.013

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93.

https://doi.org/10.1080/23273798.2015.1047459

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, *18*(3), 193–223. https://doi.org/10.1191/0267658302sr207oa

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://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. https://doi.org/10.1016/j.jml.2015.10.007

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017a). On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 635–652. https://doi.org/10.1037/xlm0000315

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017b). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition and Neuroscience*, *32*(8), 954–965. https://doi.org/10.1080/23273798.2016.1242761

Ito, A., Nguyen, H. T. T., & Knoeferle, P. (2021). Anticipatory use of verb and classifier constraints in Vietnamese-German bilinguals. Manuscript in preparation

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

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

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, *4*(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., Pugh, K. R., Shankweiler, D. P., & Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica*, *171*, 72–84. https://doi.org/10.1016/j.actpsy.2016.09.009

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S.

(2011). The time course of anticipatory constraint integration. *Cognition*, *119*(1), 23–42. https://doi.org/10.1016/j.cognition.2010.12.002

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. https://doi.org/10.1016/j.jml.2009.06.004

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, *25*(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75. https://doi.org/10.1017/S0140525X99001776

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*(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003

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. https://doi.org/10.1016/j.jml.2013.08.001

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, *19*(1), 19–35. https://doi.org/10.1017/S1366728914000881

Moors, A. (2016). Automaticity: Componential, Causal, and Mechanistic Explanations. *Annual Review of Psychology*, *67*(1), 263–287. https://doi.org/10.1146/annurev-psych-122414-033550

Neely, J. H. (1991). Semantic priming effects in visual word recognition : A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Lawrence Erlbaum Associates.

Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, *142*, 107427. https://doi.org/10.1016/j.neuropsychologia.2020.107427

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. a, Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, *7*, e33468. https://doi.org/10.7554/eLife.33468

Peters, R. E., Grüter, T., & Borovsky, A. (2018). Vocabulary size and native speaker self-identification influence flexibility in linguistic prediction among adult bilinguals. *Applied Psycholinguistics*, *39*(6), 1439–1469. https://doi.org/10.1017/S0142716418000383

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110. https://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. https://doi.org/10.1017/S0140525X12001495

Ratcliff, R., & McKoon, G. (1981). Does activation really spread? *Psychological Review*, *88*(5), 454–462. https://doi.org/10.1037/0033-295X.88.5.454

Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: A review. *Frontiers in Psychology*, *2*(NOV), 1–14. https://doi.org/10.3389/fpsyg.2011.00307

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). Oxford University Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Spivey, M. J., & Marian, V. (1999). Cross Talk Between Native and Second Languages: Partial Activation of an Irrelevant Lexicon. *Psychological Science*, *10*(3), 281–284. https://doi.org/10.1111/1467-9280.00151

Van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, *92*, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*(6), 209–216. https://doi.org/10.1016/S1364-6613(97)01070-X