Bilingualism: Language and Cognition

cambridge.org/bil

Prediction in bilingual sentence processing: How prediction differs in a later learned language from a first language

Judith Schlenter

Department of Language and Culture, UiT The Arctic University of Norway, Tromsø, Norway

Review Article

Cite this article: Schlenter J (2023). Prediction in bilingual sentence processing: How prediction differs in a later learned language from a first language. *Bilingualism: Language and Cognition* **26**, 253–267. https://doi.org/10.1017/S1366728922000736

Received: 5 January 2022
Revised: 5 September 2022
Accepted: 5 September 2022
First published online: 27 October 2022

Keywords:
prediction; sentence processing; bilingualism; visual-world eye-tracking; ERP

Address for correspondence:
Judith Schlenter
Department of Language and Culture
UiT The Arctic University of Norway
9037 Tromsø, Norway
E-mail: judith.schlenter@gmail.com

Abstract

This review provides an update on what we know about differences in prediction in a first and second language after several years of extensive research. It shows when L1/L2 differences are most likely to occur and provides an explanation as to why they occur. For example, L2 speakers may capitalize more on semantic information for prediction than L1 speakers, or possibly they do not make predictions due to differences in the weighting of cues. A different weighting of cues can be the result of prior experience from the L1 and/or the prior experience in an experiment which affects L1 and L2 processing to a different extent. Overall, prediction in L2 processing often emerges later and/or is weaker than in L1 processing. Because L2 processing is generally slower, L1/L2 differences are likely to occur at certain levels of prediction, most notably at the form level, in line with a prediction-by-production mechanism.

1. Introduction

In the past twenty years, the topic of prediction has seen a growing interest among psycholinguists after more and more studies demonstrated that listeners, signers, and readers do not just passively integrate incoming information, but use multiple information sources to predict what comes next in a sentence (for review, see e.g., Altmann & Mirković, 2009; DeLong, Troyer & Kutas, 2014; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Radošević, Malaia & Milković, 2022). When the first studies with non-native or second language (L2) speakers indicated a reduced ability to predict in the L2 as compared to the native or first language (L1), the investigation of prediction also became increasingly popular in the field of bilingualism research, as it offered the potential to shed more light on the differences in real-time processing between L1 and L2 speakers (e.g., Dussias, Valdés Kroff, Guzzardo Tamargo & Gerfen, 2013; Grüter, Lew-Williams & Fernald, 2012; Lew-Williams & Fernald, 2010; Martin, Thierry, Kuipers, Boutonnet, Foucart & Costa, 2013). As of July 2022, 41 studies have been published in peer-reviewed journals, book chapters or conference proceedings that meet the following criteria which were set for this review: (i) a time-sensitive method such as the recording of event-related potentials (ERP) or eye-tracking was used, (ii) effects of prediction were measured prior to a target constituent in line with the definition of prediction as the pre-activation of linguistic input before it is encountered by the comprehender (e.g., Huettig, 2015; Huettig, Audring & Jackendoff, 2022; Pickering & Gambi, 2018), and (iii) at least one bilingual speaker group was included (see Data Availability).

The growing number of studies on L2 prediction has changed the initial picture of prediction in L2 processing and added to the body of evidence suggesting that prediction is an integral component not only of L1 but also of L2 processing (e.g., Kaan & Grüter, 2021). However, even if L2 speakers have been shown to successfully predict upcoming linguistic information, almost all studies so far that include a between-group or between-language comparison (35 in total, 26 with adult late bilinguals) have reported at least subtle and sometimes substantial differences between L1 and L2 processing. The focus of the current review article lies on these differences (or the absence thereof). It aims to show whether and when differences are quantitative or qualitative in nature by comparing existing studies to each other and to explain the potential origins of these differences with reference to recent theoretical approaches to prediction in language processing as well as existing L2 processing accounts. Whenever I use the term 'L2 speaker,' I refer to late bilingual speakers, here defined as speakers who acquired their L2 at or after the age of six years, often under different circumstances than L1 learners or early bilingual learners. Whenever I refer to early bilinguals, I specify whether they are simultaneous or sequential bilingual speakers. Throughout, I use the term 'prediction' interchangeably with 'anticipation.'

The review is split into three subsections in an attempt to group existing studies according to the nature of the L1/L2 differences they provide evidence for: predominantly quantitative differences (2.1) and predominantly qualitative differences (2.2), as well as evidence for

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

L1/L2 differences as a result of the underlying predictive mechanism (2.3). As discussed in section 3, the distinction between quantitative and qualitative is not always clear-cut in those cases where a prediction effect in L2 processing is absent. The general discussion in section 3 summarizes the findings and points out what is special about prediction in bilingual sentence processing.

2. When prediction in L2 processing is absent or different from L1 processing

Although this review focuses on the differences in the predictive processing between groups of speakers or between the languages of a bilingual speaker, it should be acknowledged at the outset that there are many similarities and that prediction in L1 and L2 processing is affected by the same individual-level factors. In her influential review, Kaan (2014) argued that the same individual-level factors which cause variation in L1 prediction also apply to L2 prediction. As factors influencing prediction she identified (i) the frequency information stored by a speaker, which may vary particularly between L1 speakers and late learners of a language who acquire the L2 in a classroom setting and thus have less exposure; (ii) competing information; (iii) accuracy and consistency of the lexical information retrieved; (iv) task-induced processes and strategies; and (v) motivation, resources, and cognitive control. Throughout the current review, it is evident that all these factors play a role and that they are interrelated. However, with the increase in studies on prediction in bilingual sentence processing, it is now possible to pin down when L1/L2 differences are most likely to occur, and, not least thanks to more recent theoretical approaches to prediction, why. Thus, we are now in a position that allows us to better explain the nature and origin of L1/L2 differences in prediction. As shown in the first of the next three subsections, some L1/L2 differences encompass slight differences in timing or effect strength and can sometimes be so subtle that they only show up in descriptive or post-hoc analyses. Some L1/L2 differences, on the other hand, are more severe and are reflected by the complete absence of a prediction effect in the L2 or in one of several L2 groups.

2.1 Quantitative differences

Prevalent among the reported L1/L2 differences in prediction are differences in the time course of prediction. A later onset of prediction in L2 processing relative to L1 processing has been reported for the use of semantic information such as the lexical-semantics of verbs (Chun & Kaan, 2019), the semantics of classifiers (Mitsugi, 2018), negative polarity adverbs (Mitsugi, 2022), the broader semantic context (Dijkgraaf, Hartsuiker & Duyck, 2019; Ito, Pickering & Corley, 2018), gender stereotypes (Corps, Liao & Pickering, in press), the use of morphosyntactic information such as number encoded at the verb (Koch, Bulté, Housen & Godfroid, 2021) as well as discourse-level information (Kim & Grüter, 2021). In addition to differences in the time course, several studies have attested weaker effects in L2 processing (e.g., Kim & Grüter, 2021; Schlenter & Felser, 2021). In the following, I present a few examples of how L1/L2 differences surface in experimental outcomes before I discuss possible reasons for quantitative differences between L1 and L2 prediction.

So far, two studies with adult L2 speakers have reported no L1/L2 differences (Dijkgraaf, Hartsuiker & Duyck, 2017; Ito, Corly & Pickering, 2018). Both studies made use of the visual-world eye-tracking paradigm and tested the use of the lexical semantics of verbs to anticipate the upcoming object by comparing looks to the depiction of the noun in a neutral or baseline condition and a semantically restrictive condition (e.g., *The lady will find …* vs. *The lady will fold …*); an illustration is given in Figure 1. Moreover, Ito, Corley and Pickering (2018) included a cognitive load condition, in which participants had to perform an additional working memory task. L1 speakers of English and L2 speakers of English with diverse L1 backgrounds were shown a four-object display containing the target (*scarf*), a competitor from the same semantic category (*high heels*) and two distractor images. The results indicated that cognitive load hindered verb-mediated prediction in L1 and L2 speakers alike: those participants who had to perform a working memory task were slower in launching anticipatory eye movements towards the target than those without the additional task.

The two above-mentioned studies showed evidence for the successful prediction of an upcoming object based on the lexical semantics of the verb in L2 processing that emerged at the same time as in L1 processing. In line with this, a visual-world study by Corps et al. (in press) did not find an L1/L2 difference in the onset of what the authors call 'associative prediction.' However, unlike the two above-mentioned studies, the authors manipulated whether the sentences were spoken by a man or a woman and included two objects in the visual display that were compatible with the verb, but more compatible with a male or a female speaker. For example, upon listening to the sentence *I would like to wear the nice …* spoken by a male speaker, possible targets are *tie* and *dress,* but *tie* is more compatible with the agent. A four-object display contained two plausible targets (*tie*, *dress*) and two distractors (*drill*, *hairdryer*). The authors found that the integration of gender stereotypes was significantly delayed in the L2 group: English L1 speakers in Corps, Brooke and Pickering (2022) predicted the verb-compatible target (*tie* in comparison to *drill*) 519 ms after verb onset and the verb- and agent-compatible target (*tie* in comparison to *dress*) 641 ms after verb onset. Highly proficient L2 speakers of English with diverse L1 backgrounds predicted the verb-compatible target 527 ms after verb onset and the verb- and agent-compatible target only from 957 ms after verb onset. Thus, when also the perspective of the speaker had to be taken into consideration for the prediction of an upcoming object, L2 speakers needed more time than L1 speakers.

In the following eye-tracking study to be discussed, time course differences were found to be coupled with less certainty in L2 processing. Mitsugi (2018) tested the use of numeral classifiers in Japanese to predict the upcoming noun. Nouns in Japanese fall into different classes and, when quantified, are preceded by a classifier associated with that class. This mapping can be semantically prototypical – for example, the classifier *-hon* is typically associated with long and string-like objects, as in example (1) from the study, but it can be also used more metaphorically (e.g., for phone calls). This study was restricted to prototypical classifier-noun pairings.

(1) soredewa san-**hon**-no kasa-o
Well-then three-CL-GEN umbrella-ACC
kurikku-shite-kudasai
Click-do-please
'Well then, please click on the three umbrellas.'

Prediction was investigated by comparing looks to the target image between informative and uninformative trials. In

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

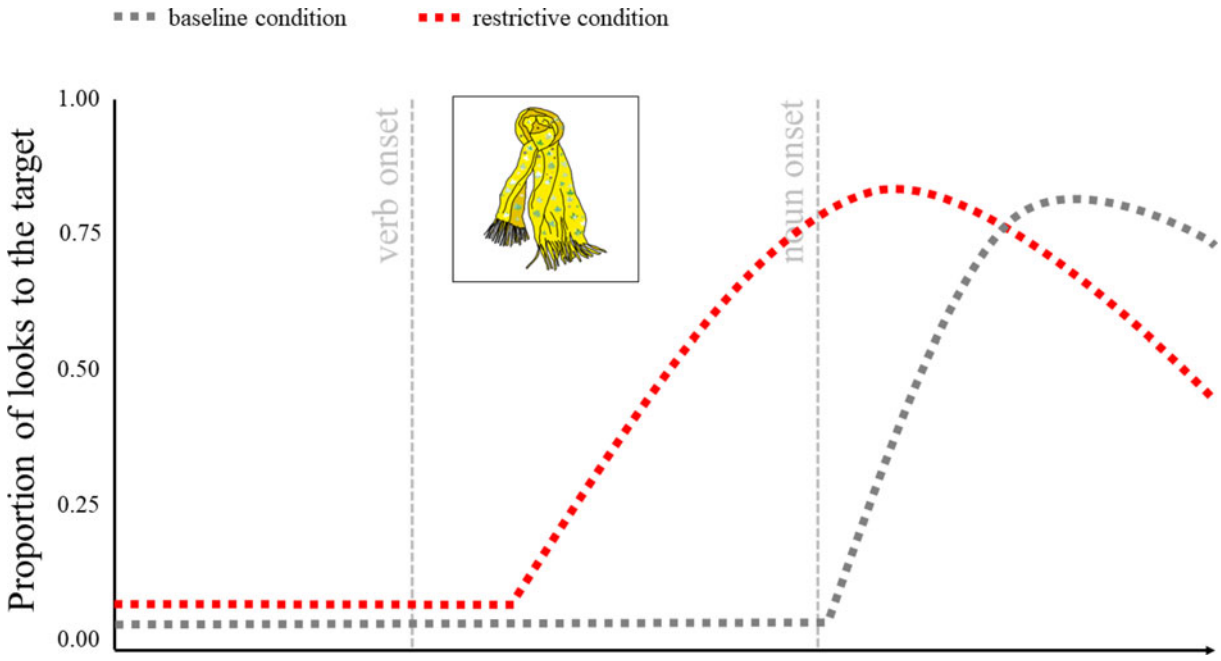


Fig. 1. An illustration of an experimental design in which looks to a target image (taken from the MultiPic database, Duñabeitia et al., 2018) are compared between a baseline condition and a condition with a predictive cue. In this example, the predictive cue is the verb *fold* that restricts the domain of subsequent reference to objects that are foldable. If a semantically neutral verb is heard (e.g., *find*), participants should look towards the target only after noun onset.

informative trials, the two images that were presented belonged to different noun classes, while in uninformative trials both objects belonged to the same noun class. Both L1 speakers of Japanese and intermediate proficient English L1–Japanese L2 speakers were more likely to fixate on the target image prior to noun onset in informative trials. However, separate analyses per group revealed differences in the time course: the growth curve analyses that were conducted showed that target fixation quickly increased in the L1 group and then leveled off. In the L2 group, the increase was more gradual, and the effect of condition persisted during the noun region. Thus, while the classifier cue was immediately effective in the L1 group, it was less effective in the L2 group, and the prediction effect developed more slowly. Moreover, the L2 group demonstrated fewer target fixations in the uninformative condition than the L1 group. Overall, Mitsugi (2018) noticed that the L2 speakers appeared less certain than L1 speakers, and their gaze pattern was characterized by wandering around, an observation also made by Corps et al. (in press). Results from analyses including L2 proficiency were inconclusive: L2 proficiency did not mediate prediction, but it was found to mediate the integration of the noun in uninformative trials. The more proficient the L2 speakers were, the faster the visual attention was shifted to the target image.

The studies discussed here are examples of typical patterns seen in L2 prediction studies that employ the visual-world paradigm: the point in time when looks to a target image and other images diverge after the onset of a cue (e.g., a classifier) is later in an L2 as compared to an L1 group. Often the L2 group is found to show fewer looks to the target images and to be more distracted by other 'non-target' images than the L1 group. Another common observation is the lack of a correlation between prediction and L2 proficiency (for discussion, see Kaan & Grüter, 2021) and, for L1 and L2 processing alike, a correlation between prediction and domain-general executive resources (for discussion, see Ryskin, Levy & Fedorenko, 2020). Although these factors can play a role in prediction, they may not always do so, indicating that other factors are probably more important for the prediction of upcoming linguistic input. In the following, I aim to show how the more quantitative L1/L2 differences in prediction can be accounted for by task-induced processes, more precisely, the complexity of the experimental stimuli and time constraints. Before that, I describe a factor – or rather a combination of interrelated factors – that might be responsible for a large share of variability within and between groups in experimental studies.

Lexical processing

As highlighted by Kaan (2014), frequency information as well as the accuracy and consistency of the lexical information retrieved can affect prediction – for example, Kaan referred to studies demonstrating that L1 speakers' prediction can be dependent on vocabulary size (Borovsky, Elman & Fernald, 2012), literacy skills, or reading and writing experience (Mani & Huettig, 2014; Mishra, Singh, Pandey & Huettig, 2012); for a recent overview, see Huettig and Pickering (2019). Further evidence for a connection between lexical representation and prediction in L1 and L2 speakers was provided by Hopp (2013). In an eye-tracking task, Hopp examined whether L1 and L2 speakers of German were faster in shifting their gaze to a target object in trials in which the grammatical gender cue on the definite article was informative regarding the upcoming noun. The results revealed that English L1–German L2 speakers who consistently assigned the correct lexical gender in a production task were also more likely to predictively process syntactic gender agreement. Thus, some L1/L2 differences in prediction may have their origin at the level of lexical processing. If a language comprehender has a stable lexical representation of a word – including, for example, the lexical gender for nouns or the subcategorization information for verbs – the more easily it should be retrieved and used for prediction. This is in line with the Lexical Bottleneck Hypothesis (Hopp, 2018), according to which difficulties with lexical processing affect later syntactic

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

processing such as the processing of gender agreement. In the next section, I show how L1/L2 differences in lexical processing can interact with task-induced processes.

Complexity

We saw earlier that eye-tracking studies on semantic prediction reported successful prediction in L2 processing. In general, we can posit that L2 speakers are able to use verb semantics to anticipate the target object when visually presented with several objects to choose from. A mechanism that might have aided prediction in such experimental designs is association. More recent theoretical approaches to prediction discuss different forms of priming as mechanisms subserving linguistic prediction, among them priming through semantic association (e.g., Huettig, 2015; Huettig et al., 2022; Pickering & Gambi, 2018). Although described as less effective than other mechanisms of prediction, 'prediction-by-association' likely still facilitates language comprehension (e.g., Pickering & Gambi, 2018). However, semantic associations alone may not suffice when other sources of information need to be integrated into the parse, such as the perspective of the speaker in the study by Corps et al. (in press). In line with this idea, the authors interpret the finding that agent- and verb-consistent prediction (predicting *tie* rather than *dress* for a male speaker) emerged later than verb-consistent prediction (predicting *tie* rather than *drill* after encountering the verb *wear*) as evidence for a second stage of prediction. Thus, adding complexity can lead to quantitative differences in L1/L2 prediction. This has also been shown by Chun and Kaan (2019), who tested the use of verb semantics in a comparison between semantically restrictive (*open*) and neutral verbs (*get*), using complex sentences (e.g., *I know the friend of the dancer that will open/get the present*) and visual scenes with multiple objects. In this study, it took the L2 group 180 ms longer than the L1 group to anticipate the target image.[1] Prediction in the L2 group was not mediated by L2 proficiency or working memory.

To wrap up, one possible explanation for the variability in study outcomes is a difference in complexity: if complexity increases, be it linguistically or non-linguistically, differences in prediction between L1 and L2 processing are more likely to occur. Complexity, and with that uncertainty on the part of the L2 speakers, may increase even more when the visual display not only contains multiple images, but when it contains competitor images – as, for instance, shown by Peters, Grüter and Borovsky (2018) (but cf. Ito, Corley & Pickering, 2018, who found that a semantic competitor image did not receive more fixations than two distractor images). Peters et al. (2018) presented a large heterogeneous group of bilingual speakers with sentences such as *The pirate chases the ship* while the visual display, in addition to the target image (*ship*), contained the image of an agent-related object (*treasure*), an action-related object (*cat*), and an unrelated object (*bone*). The authors found that the consideration of competitor images was more pronounced in bilingual speakers who had a smaller vocabulary and/or speakers who indicated themselves as non-native speakers of English: low-vocabulary bilingual speakers and self-indicated L2 speakers were more likely to consider the image of a cat upon listening to the action verb than high-vocabulary speakers and speakers who considered themselves as native speakers. In a study on the predictive use of placement verbs in Dutch by van Bergen and Flecken (2017), the visual display showed four images (e.g., four objects on a table: a ball, a lying bottle, a standing bottle, a cake) that would be more or less likely targets depending on the information just given. For example, when listening to *De jongen zette kort geleden een fles op de tafel* ('The boy put recently a bottle on the table') the placement verb *zetten* provides a cue towards the orientation of the object – that is, standing – so both the bottle in standing position and the cake in the display are possible targets. The results showed that Dutch L1 and advanced German L1–Dutch L2 speakers successfully predicted the object's position upon encountering the verb and fixated more on the standing objects than the two other objects in the display. However, only the L1 group immediately fixated the respective target image during the verb-object integration window (i.e., bottle standing on table). The German L1–Dutch L2 speakers, on the other hand, did not readily integrate the noun information with prior verb information, and initially also considered the position competitor image (i.e., bottle lying on table). This likely indicates that predicting upcoming linguistic input affected L1 and L2 processing to a different extent: integration of the noun was readily facilitated in the L1 but not in the L2 group. We may conclude that even advanced L2 speakers can be affected by the simultaneous presentation of competitor images because predictions and overall processing are less certain in an L2 than in an L1.

Methodological factors might also have contributed to L1/L2 differences in the following two studies. Dijkgraaf et al. (2019) compared looks to a target image and three distractor images while manipulating whether the target image corresponded to the target word or a semantic competitor. For example, participants would listen to a sentence such as *Her baby doesn't like drinking from …* and would either be presented with the image of a bottle or the image of a glass – that is, a noun that is compatible with the verb, but is a less likely continuation. In a within-group comparison, the authors found that Dutch L1–English L2 speakers' prediction was slower in their L2 English as compared to their L1 Dutch for the semantic competitor condition. Additionally, the results showed that the strength of the association between target and semantic competitor, as measured by semantic distance scores, influenced prediction in both the L1 and L2, yet less in the L2. There was no effect of proficiency. The authors took their findings as an indication for a slower and weaker spreading of semantic activation in L2 as compared to L1 processing. They also noted that their sentences were more complex than in previous studies and required participants to use more than just verb semantic information. In addition, and this might be another crucial aspect, the images became visible only 500 ms prior to target onset. Less efficient semantic prediction together with methodological factors may also account for the L1/L2 differences observed by Schlenter and Felser (2021, Experiment 1). The authors found that both L1 speakers of German and Russian L1–German L2 speakers used the lexical semantics of the verb to predict the animacy of the direct object. After verbs that semantically select an animate object (e.g., *The woman feeds the black …*), participants across groups were more likely to fixate on the image of the animate entity *cat* than the image of the inanimate entity *shirt*. After verbs that semantically select an inanimate object (e.g., *The woman irons the black …*), looks towards the inanimate entity increased. However, separate analyses showed that the prediction effect was not significant in the L2 group for the animate-biasing condition, indicating less certainty for those verbs in the study that semantically

[1]Note that due to the absence of comprehension questions in this study, Chun and Kaan (2019) could not rule out that participants paid attention only to the verb without fully understanding the sentences; evidence for successful prediction in a design with comprehension questions and a comparison between simple and complex sentences in L2 speakers is provided by Chun, Chen, Liu and Chan (2021).

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

select an animate object. When compared to other studies, the prediction region in this experiment was rather short with around 500 ms given for prediction. Ito, Corley and Pickering (2018), for example, point out that their sentences were spoken slowly and with pauses, giving participants an average duration of 1959 ms from verb onset until target noun onset. The findings that were just described are compatible with differences in lexical processing between L1 and L2 speakers. Moreover, it is very likely that methodological factors enhanced the L1/L2 differences that were observed. With this, we turn to another important methodological factor in prediction studies, the timing.

Time constraints

Not only the accuracy and consistency of lexical information, but also the speed of lexical access can mediate prediction in L1 and L2 processing. In the study by Hopp (2013) discussed above, the speed of lexical access was measured by a control condition (reaction time after lexical cue, e.g., *Where do you see two* [*adjective*][*noun*]*?*). Participants across groups who needed more time in accessing lexical information showed weaker effects of prediction based on gender. Lexical processing is likely to take longer especially in bilingual sentence processing, not the least due to the co-activation of both languages of the bilingual (Hopp, 2018). Thus, the activation of competing information may further slow down L2 prediction. Together with experimental design choices (e.g., complexity of the visual display and preview time) as well as properties of the materials that affect the time given to predict, this may lead to the complete absence of a prediction effect in L2 processing, as demonstrated in the following study by Fernandez, Engelhardt, Patarroyo and Allen (2020). The time given for prediction depends on the inclusion of intervening material between cue and target constituent and/or the presentation rate of the stimuli, as in the following study.

Fernandez and colleagues investigated whether listeners predict the noun corresponding to the patient/theme of an event, thereby actively forming a filler-gap dependency. After a short context followed by a *wh*-question or point-to probe (e.g., *Who$_i$ did the wolf attack t_i near the cave?*), listeners of English were expected to preferentially fixate on the picture of the patient as compared to the agent in a four-object display upon encountering the verb. Fixations were analyzed in a one second time window starting 200 ms after verb onset. Crucially, the authors' primary focus was to find out how different age and language groups were affected by the speech rate. To this end, they compared three groups of English speakers: adult L1 speakers (age range 18–24), older L1 speakers (age range 40–75), and late L2 learners of English with L1 German who were university students. The sentences were provided at speech rates of 3.5 syllables per second and 4.5, 5.5 and 6 syllables per second. While young L1 speakers were able to predict at rates up to 5.5 syllables per second, older L1 speakers only predicted at rates of 3.5 and 4.5 syllables per second. Furthermore, adult L2 speakers only showed anticipatory eye movements toward the target at a speech rate of 3.5 syllables per second. The results thus indicate that L2 speakers need more time than L1 speakers of the same age and older to generate predictions. In other words, L2 speakers need to be given sufficient time or they may show no effect of prediction, as in the study by Fernandez et al. (2020). Time constraints were also evident in L1 processing but turned out to be particularly critical in L2 processing.

2.2 Qualitative differences

In the following, the focus lies on studies that report a lack of prediction in L2 processing while prediction is evident in L1 processing. Note that absence of prediction does not necessarily imply that L2 speakers did not predict at all, but it can also refer to studies where they did not predict in all experimental conditions.

One information source that has been shown to be notoriously difficult to use predictively in L2 processing is the morphological marking of case. This was first attested by Hopp (2015). There, L2 speakers of German relied exclusively on verb semantics and the linear order of words to anticipate the final argument in subject- and object-initial transitive sentences – that is, irrespective of the case marked on the first NP. For example, in an object-initial German sentence like *Den Wolf tötet sogleich …* ('The-ACC wolf kills soon …'), L2 speakers moved their eyes towards the depicted *deer* instead of to the depicted *hunter*. L1 speakers, in contrast, integrated both morphosyntactic and lexical-semantic information to anticipate a plausible agent or patient. Proficiency in German could not account for the absence of prediction based on morphological case in the L2 group (but see Henry, Jackson & Hopp, 2022, who observed more target fixations at higher proficiency levels). In an eye-tracking study by Mitsugi and MacWhinney (2016) in the verb-final language Japanese, prediction based on morphological case was, again, only evident in an L1 group, and absent for a group of intermediate-level English L1–Japanese L2 speakers. In the design used by the authors, looks to a plausible theme argument (*test*) were compared between a canonical argument order (serious student-NOM strict teacher-DAT), a scrambled order (strict teacher-DAT serious student-NOM) and a baseline condition. In the baseline condition, only a monotransitive verb could follow and no additional argument (serious student-NOM strict teacher-ACC). L1 speakers were found to anticipate the theme argument (test-ACC) in both ditransitive verb conditions. Only after noun onset, L2 speakers looked more at the theme for the ditransitive verb conditions relative to the baseline condition. In a subsequent study by Mitsugi (2017), testing sentences with monotransitive verbs such as *hit*, L1 speakers of Japanese predicted the voice of the sentence-final verb upon encountering the second NP. Case marking at the second NP signaled whether it was a dative argument, which can only be followed by a verb with passive morphology, or an accusative argument and thus should be assigned the patient role in an active sentence. The participants were presented with two scenes, one showing the target event (e.g., woman hitting man) and another showing the same event with role reversal (man hitting woman). Prior to the verb, the L1 group was more likely to fixate on the target than the competitor scene. Again, prediction was absent in a group of English L1–Japanese L2 speakers. Only after encountering the verb, L2 speakers were more likely to fixate on the target scene.

While the results of these studies might indicate that prediction based on a morphosyntactic cue such as case marking is impossible in L2 processing, the above findings might also be explained by the fact that all three studies included L1 speakers of English, a language with a rigid word order and no case marking on full NPs. Thus, there could be some cross-linguistic influence involved. Some indication for cross-linguistic influence also comes from a study on Korean by Frenck-Mestre, Kim, Choo, Ghio, Herschensohn and Koh (2020) that included two groups of L2 speakers, Kazakh L1 and French L1. Although neither of the two L2 groups showed a predictive use of case in Korean, unlike the Korean L1 group, the two groups differed in their final sentence interpretation, with the L1 Kazakh group showing a more native like eye-gaze pattern and accuracy than the L1

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

French group. French has no case marking and a less flexible word order than Korean and Kazakh. Since the L2 speakers had been studying Korean for three semesters only and thus were probably not advanced L2 speakers, the results leave open the question of whether, at higher proficiency levels, L2 speakers whose L1 shares properties with the L2 (case marking, flexible word order) can use case predictively. Recent findings from Schlenter and Felser (2021, Experiment 2) indicate that highly proficient Russian L1–German L2 speakers were sensitive to morphological case for word order manipulations after a ditransitive verb and used it predictively. Like German, Russian is a case-marking language with flexible word order. Moreover, in this experiment, case was not only marked on the pre-nominal article but, especially since Russian lacks an article system, also on a pre-nominal adjective (e.g., *Der Gärtner gibt der/die blühenden/e Pflanze …* 'the gardener gives the-DAT/ACC flowering-DAT/ACC plant …'). However, this study lacked an additional L2 group with no case marking in the L1, so it is unclear whether the predictive use of morphological case in this study resulted from the L2 speakers' familiarity with case, the materials, or a combination of both.

Taken together, almost all studies that investigated the predictive use of case marking found an L1/L2 difference with successful prediction in an L1 group but no prediction in an L2 group. However, it cannot be ruled out that a lack of familiarity with morphological case and flexible word order was the main factor driving the L1/L2 difference. In fact, several L1/L2 differences reported in the literature might have been caused by cross-linguistic differences, e.g., the limited evidence for syntactic prediction in a Dutch L1–English L2 group in Kaan, Kirkham and Wijnen (2016). In this ERP study, L1 speakers of English used the syntactic context to predict an elliptical structure in sentences such as (2-a).

(2) a. Although Peter met John's surgeon, he did not meet Max's […]
b. *Although the surgeon met John, he did not meet Max's […].

To analyze prediction, ellipsis conditions such as (2-a) were compared to non-ellipsis conditions (2-b). In a time-window between 500 and 700 ms after the onset of the possessive in the second clause, the L1 group showed an increased positivity in reaction to non-ellipsis conditions, while the results for the L2 group remained inconclusive. Since the L2 speakers' L1 Dutch does not allow an ellipsis after possessive proper names, prediction in the L2 group might have been hindered by cross-linguistic differences. The following study with early bilingual children by Meir and colleagues lends further support to the assumption that the predictive use of morphological case can be influenced by the degree of cross-linguistic overlap. At the same time, this study shows that it is probably not just a question of whether a predictive cue exists in the bilingual speaker's other language, but whether it is similarly weighted.

In an eye-tracking study, Meir, Parshina and Sekerina (2020) compared Russian-Hebrew bilingual children, who lived in Israel and spoke Russian as their heritage language (age of acquisition for Hebrew: 0–4) to a monolingual Hebrew control group and a monolingual Russian control group, in order to investigate whether monolinguals and bilinguals differ in their predictive use of morphological case (age at testing: 4–8). While monolingual Hebrew-speaking children, unlike monolingual Russian-speaking children, were insensitive towards case and did not use it predictively, bilingual Russian-Hebrew-speaking children used the case marker on the first NP to anticipate the upcoming argument's thematic role, agent or patient, in both Russian and Hebrew. The authors argue that a weak cue in one language, here Hebrew, was reinforced by a stronger cue in the other language. Case morphology has been found to be a strong and reliable cue to identify thematic roles in Russian. Currently, the age at which speakers of Hebrew start using case marking predictively is unknown. Thus, it remains to be addressed whether the bilingual children had an advantage over the monolingual Hebrew-speaking children. For Russian, the bilinguals' use of case was slightly delayed as compared to the monolingual group. Altogether, the group comparisons yielded evidence for positive transfer from Russian to Hebrew in the bilingual child group.

To conclude this section, while the predictive use of a cue such as morphological case seems to be difficult in adult L2 processing, as indicated by its absence in L2 groups with no familiarity with case from their L1 or less proficient L2 speakers (Frenck-Mestre et al., 2019; Hopp, 2015; Mitsugi, 2017; Mitsugi & MacWhinney, 2016), it has been found to be achieved under what I will henceforth call 'favorable conditions' (Schlenter & Felser, 2021; see also Henry et al., 2022). In the next sections, we consider what these favorable conditions are.

Cross-linguistic influence

Oftentimes, there are multiple explanations as to why there is a difference in the predictive processing between an L1 and L2 group. Thus, an L1/L2 difference is not necessarily more informative as regards the underlying cause than no L1/L2 difference. However, if an effect of prediction shows up in one L2 group but not another L2 group whose L1 differs in a specific characteristic from the L1 of the first group, then we have good reason to infer that one cause was the L1/L2 overlap or lack thereof. As outlined at the beginning of this review, we know by now that L1 and L2 prediction is affected by the same individual-level factors (Kaan, 2014). Cross-linguistic influence (CLI), however, is unique to bilingual sentence processing. If we want to test L2 prediction under favorable conditions, then we need to include an L2 group whose L1 encodes the information that is predicted and the information that serves as a predictive cue (e.g., morphological case on NPs). Moreover, if we want to know whether and how L2 prediction is affected by CLI, we should ideally include more than one L2 group (and match these groups in terms of proficiency, age of L2 onset, etc.), since otherwise, as highlighted by Lago, Mosca and Stutter Garcia (2021), the application of general-purpose processing mechanisms by L2 speakers may be falsely attributed to L1 influence. Cross-linguistic influence can take different forms. The absence or different realization of certain features in the languages a bilingual speaks can impede prediction in the L2. Alternatively, if a feature exists in both languages, L2 prediction can be facilitated (for a systematic review, see Foucart, 2021).

Compelling evidence for CLI comes from an eye-tracking study on semantic prediction by van Bergen and Flecken (2017). The authors investigated whether advanced L2 learners of Dutch predict the positioning of an object after encountering a placement verb. Whereas German, like Dutch, specifies the object position in placement verbs (Dutch: *zetten*, German: *stellen* 'put.STAND' and Dutch: *leggen*, German: *legen* 'put.LIE'), it is not specified in English and French, which use one general placement verb (English: *put*, French: *mettre*). The Dutch L1 and German L1–Dutch L2 speakers in the study demonstrated anticipation

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

of the objects' perceptual features – that is, standing or lying. In contrast, a group of English L1–Dutch L2 and a group of French L1–Dutch L2 speakers did not. The verb effect only became significant in the time window that ranged from 500 ms until 1000 ms after noun onset. Thus, only after the information became available, these two groups fixated on the respective target image. This latter finding confirmed that they understood the meaning of the placement verbs and it supports the authors' conclusion that the presence/absence of prediction was the result of positive/negative CLI.

Further evidence for CLI comes from an ERP study by Alemán Bañón and Martin (2021), who investigated whether prediction at the level of syntax was influenced by cross-linguistic overlap between the L1 and L2. To that end, the authors tested the prediction of a genitive construction in an L1 English control group and two groups of L2 speakers, Swedish L1 and Spanish L1. While Swedish and English have possessive pronouns that mark the referent's natural gender (*hans/hennes* 'his/her'), Spanish possessive pronouns only mark the gender of the possessed noun. Participants read sentences, such as the ones presented in (3), that included either an expected (3-a) or unexpected possessive pronoun (3-b) – based on the prior discourse context – followed by a kinship noun.

(3) Julia and Albert have joined a meditation group. Julia's niece dislikes being quiet and hates meditation. However, Albert's sister really enjoys silence and loves meditation. In your opinion, which of the two will they invite to meditation, Julia's niece or Albert's sister.
   a. In my opinion, it is his sister that they will invite.
   b. In my opinion, it is her niece that they will invite.

Unexpected possessive pronouns induced an N400 for English L1 and Swedish L1–English L2 speakers alike. In contrast, Spanish L1–English L2 speakers showed a P600. The authors explained their findings in terms of qualitatively different predictions among the group of L2 learners whose L1 does not encode the natural gender of the possessor. Thus, in this study, a different encoding of features did not lead to the absence of a prediction effect in the L2 but resulted in a deviant ERP response.

Indication for CLI is also provided by the following study including two experimental conditions that differed in L1/L2 overlap. In a study on predictive gender agreement, Hopp and Lemmerth (2018) tested Russian L1–German L2 speakers with high-intermediate to advanced proficiency in German. Like German, Russian is a gender-marking language. Unlike German, Russian has no articles, but marks gender post-nominally on suffixes or, like German, pre-nominally on adjectives. Examples from the study are provided in (4).

(4) a. *Wo ist der/die/das gelbe* [noun]?
   Where is the[masc.]/[fem.]/[neut.] yellow [noun]?
   b. *Wo ist ein kleiner/s gelber/s* [noun]?
   Where is a small[masc.]/[neut.] yellow[masc.]/[neut.] [noun]?

While the advanced group showed a nativelike gaze pattern in the eye-tracking task, the high-intermediate group only did so for the adjective condition (4-b). When gender was marked on the article (4-a), the high-intermediate group anticipated the upcoming noun only if its lexical gender was the same in German and Russian, e.g., both feminine. The results thus indicate that, at lower proficiency levels, lexical and syntactic congruency between the languages affected L2 prediction. Recent studies with bilingual children add to the body of evidence suggesting that L1/L2 overlap influences prediction, but also point towards differences between early and late bilinguals (Bosch, Chailleux, Yee, Guasti & Arosio, 2022; Lemmerth & Hopp, 2019; for review, see Karaca, Brouwer, Unsworth & Huettig, 2021). For example, in a comparison between monolingual German, simultaneous Russian–German bilingual and sequential Russian–German bilingual children, Lemmerth and Hopp (2019) found that the sequential bilingual children did not predict if gender was incongruent between languages, irrespective of their proficiency.

L1/L2 differences in cue weighting

The experimental findings discussed in this section relate to differences in cue weighting between L1 and L2 groups as they have been proposed in different psycholinguistic models. Before we turn to experimental findings that were not yet discussed in this review, we return to the example for L1/L2 differences in prediction when morphological case provided a cue which thematic role will follow: the explanation that English L1 speakers do not use the morphological marking of case in their L2 because case marking on NPs is non-existent in their L1 finds further support by their alternative strategy, as evidenced first in Hopp (2015). To recap, these speakers used the lexical semantics of the verb together with the linear order of words to predict the postverbal argument – that is, instead of waiting for the second argument to appear, they predicted a patient (*deer*) after encountering the first argument (*wolf*) and the transitive verb (*kill*) although accusative marking on the first argument should have ruled out an SVO interpretation. Thus, the L2 speakers over-relied on the other cues in the sentence. An over-reliance on semantics in an L2 group was also visible in an eye-tracking study by Grüter, Lau and Ling (2020). Here, English L1–Mandarin Chinese L2 speakers relied more on the semantics of classifiers and less on class membership. For L1 speakers, in contrast, form class was the primary cue. The authors were able to detect this difference because they included non-prototypical classifier-noun pairings such as *tiáo-gǒu* (CL-LONG-THIN-dog), where the noun *dog* is a less prototypical class member than, for example, the noun *rope*. The classifier *tiáo* typically appears together with long, slender, flexible objects.

The findings by Hopp (2015) can be accounted for by the Competition Model (Bates & MacWhinney, 1989). The model was designed to quantify the ways in which distributional properties of the input control the learning and processing of a language. Its underlying assumption is that the language learner needs to detect certain cues and their reliability and availability, which determine their strength. Pre-verbal positioning has been identified as a strong cue for subject/agent identification in English (MacWhinney, 2008), and it is likely transferred to the processing of L2 German. Meir et al. (2020) also interpret the differences between mono- and bilingual children found in their study in the context of the Competition Model. To recap, here a group of monolingual children did not use case marking predictively in Hebrew, while a group of Russian-Hebrew bilingual children did. The findings by Hopp (2015) as well as the findings by Grüter et al. (2020) can also be accounted for by the Shallow Structure Hypothesis (Clahsen & Felser, 2006), as they show a greater reliance on surface-level and semantic information in an L2 as compared to an L1 group. The

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

Shallow Structure Hypothesis (SSH) posits that L2 speakers are more strongly guided by nongrammatical information than L1 speakers during processing and, more recently, that "processing reflexes of grammatical information might be delayed in L2 compared to L1 processing, and/or in comparison to processing reflexes of nongrammatical information" (Clahsen & Felser, 2018, p. 11).

Differences in cue weighting might also account for the findings of the following study investigating discourse-level prediction. There, L2 speakers did not over-rely on a semantic cue as in the studies above, but they did not integrate a semantic and a grammatical cue, unlike an L1 group. In Grüter, Takeda, Rohde and Schafer (2018) and Grüter and Rohde (2021), participants had to integrate the lexical semantics of the verb and grammatical aspect to predict upcoming reference. For transfer-of-possession verbs such as in (5), more references to the goal argument were expected after completed events. In visual-world eye-tracking, this should be reflected by more looks to the image of the goal argument (*Emma*) prior to the onset of the disambiguating pronoun. This expectation was borne out for a group of L1 speakers, who were found to look more to the goal argument after perfective-marked verbs (*gave*) than after imperfective-marked verbs (*was giving*) during the intersentential pause (Grüter et al., 2018).

(5) $\text{Patrick}_{\text{SOURCE}}$ gave/was giving $\text{Emma}_{\text{GOAL}}$ a bottle of nice wine. He/She obviously knew about fancy food and drink.

In contrast, Grüter and Rohde (2021) found that grammatical aspect did not modulate the prediction of upcoming reference in L2 speakers of English. The L2 speakers with varying L1s showed no difference between conditions in real-time processing, despite their native-like performance in an additional offline task. L2 proficiency could not explain the findings.

The above-mentioned findings by Grüter and Rohde (2021) can be explained by the SSH, as L2 speakers showed no use of grammatical information for prediction. Alternatively, the L1/L2 difference in the study by Grüter and Rohde (2021) can be explained by the model proposed by Cunnings (2017), according to which L2 speakers are more susceptible to interference during memory retrieval operations than L1 speakers. As shown in (5), towards the end of the first clause, participants in the study had to retrieve information provided at the verb to interpret (upcoming) reference. Thus, L1/L2 differences in this study could have been the result of how cues to memory retrieval were weighted.

To sum up, several studies showing L1/L2 differences in prediction can be explained by differences in cue weighting. Crucial for the purpose of the current review, one conclusion we can draw from the finding that L1 and L2 speakers weigh cues differently is that more qualitative L1/L2 differences in prediction are likely to occur when a predictive cue does not exist in the L2 speakers' L1/is less reliable in the L1 (classifiers associated with noun classes, Grüter et al., 2020; morphological case on NPs, Hopp, 2015) and/or when different sources of information need to be integrated for prediction (Grüter & Rohde, 2021; Hopp, 2015).

L1/L2 differences in the utility of prediction

Previously, we saw that whether a cue is used for prediction may depend on the weighting of this cue in the languages of a bilingual speaker – for example, morphological case marking on NPs in German should be a less reliable cue for the prediction of an upcoming thematic role for L1 speakers of English than for L1 speakers of Russian. Now, we turn to the reliability of a predictive cue within an experiment. It is known that even L1 speakers may not predict, for example, the gender of a noun when they are exposed to inconsistent gender assignment during an experiment (Hopp, 2016). Moreover, L1 prediction can be facilitated through the presence of multiple cues, e.g., a morphological case cue as well as a prosodic cue (Henry, Hopp & Jackson, 2017). In the following, I illustrate how the presentation of inconsistent cues might have hindered prediction in L2 processing.

The following studies examined the predictive use of prosodic cues in L2 processing, leading to mixed results. Results from an eye-tracking study by Perdomo and Kaan (2021) showed that L1 speakers of English, but not Mandarin Chinese L1–English L2 speakers, immediately looked more towards the same object (Benjamin's cake if cake was heard before) when the second proper name in the critical sentence carried a contrastive pitch accent (6-b) than when it did not (6-a). Both L1 and L2 speakers displayed an overall preference towards the different or new object (e.g., *cake* if ice cream was heard before).

(6) a. We ate Angelas' ice cream/cake but saved Benjamin's cake in the fridge.
b. We ate Angelas' ice cream/cake but saved BENjamin's cake in the fridge.

In the L2 group, an effect of accent only became significant 200–400 ms after the onset of the noun. Indication of a modulation by working memory was absent, as was a clear indication for a modulation by L2 proficiency. The findings align with those from a study by Foltz (2021b) that tested the use of prosodic information in L1 and L2 processing within the same group of bilingual speakers. Foltz found that German L1–English L2 speakers' use of contrastive pitch accent to anticipate the upcoming noun was restricted to their L1. However, as in Perdomo and Kaan (2021), the prosodic cues were only available in two of the four conditions and were inconsistent in half of the trials; see example (7).

(7) Click on the red duck. Click on the green/GREEN … (duck/banana).

It could be the inconsistency of the predictive cue that has impacted the results, as indicated by the results from a second experiment in Foltz (2021b). Here, the L2 speakers used the cue in a predictive fashion after exposure to 24 trials with consistent prosodic cues. Further empirical support for the successful integration of prosodic information comes from a study by Henry et al. (2022). The authors did not only test whether English L1–German L2 speakers integrate case marking and verb semantics to predict the postverbal argument, but also manipulated the prosody of the sentence-initial NP. They found that L2 speakers, like L1 speakers (Henry et al., 2017), were more likely to fixate on the target image when the sentence-initial NP additionally received contrastive accent and thus was also prosodically marked.

The inclusion of inconsistent information might also account for some quantitative differences reported in the literature. For example, Kim and Grüter (2021) found a difference in timing between English L1 speakers and Korean L1–English L2 speakers in using implicit causality information to predict upcoming reference. Moreover, the prediction effect in the L2 group not only had a later onset but was also weaker in the L2 relative to the L1 group. In this study, half of the experimental trials included sentence continuations that were incongruent with the implicit causality

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

bias of the verb. In example (8) from the study, the verb *bother* should exert a bias towards an explanation including Patrick as the causer of the event. However, as in the example, in half of the trials the continuation was plausible but not consistent with the verb's bias, which might have rendered the verb's implicit causality bias less reliable.

(8) Patrick and Curtis were solving math problems in class. Patrick bothered Curtis every few minutes because he was the smartest kid in class.

In contrast, a high and noticeable reliability of a cue may have a facilitative effect. For example, Koch et al. (2021) noted that the Dutch L1–German L2 speakers in their study indicated in a post-experiment interview that they became aware of the cue: here, number marking on the verb. Thus, in addition to familiarity with subject-verb agreement from their L1 Dutch, becoming consciously aware that there was a cue might have aided their prediction.

Several researchers have argued that the strength of prediction is influenced by statistical contingencies between stimuli, and that the parser adapts to the input to maximize the utility of prediction and reduce the costs of potential prediction errors (e.g., Kuperberg & Jaeger, 2016). More recently, studies have shown evidence of adaptation in L2 predictive processing, but also hint at L1/L2 differences in adaptation (e.g., Foltz, 2021a, 2021b; Hopp, 2021). Consequently, the utility of prediction has been discussed as another factor with the potential to explain the inter- and intra-individual variability of prediction (e.g., Grüter & Rohde, 2021; Kaan & Grüter, 2021). Grüter and Rohde (2021) proposed that L2 speakers may weigh the costs of prediction to a different extent than L1 speakers and that in some instances the benefits of prediction do not outweigh its costs. Hence, if a cue reliably predicts upcoming linguistic input, it is likely that this cue is used in an anticipatory fashion. If not, L2 speakers may wait for upcoming linguistic information to appear to avoid costly re-analysis. As highlighted in previous sections, the predictive use of certain information sources in bilingual sentence processing can be more complicated when languages differ in the weighting that they assign to them or when information is not or differently encoded. This may go unnoticed unless researchers specifically test for (subtle) influences from other languages.[2]

To conclude, favorable conditions for L2 prediction include consistent and thus reliable predictive cues. This entails that also the experimental design and materials can influence whether language comprehenders and particularly L2 comprehenders predict upcoming linguistic information.

2.3 L1/L2 differences at the level of prediction

Up to now, this review has mostly focused on the use of a predictive cue or a combination of cues in an experiment. In the following, I review the evidence showing that L1/L2 differences can be tied to the level of prediction. Even under favorable conditions, L1/L2 differences are likely to occur at the level of phonological form and less likely to occur at the syntactic and conceptual level (see also Ito & Pickering, 2021).

In an ERP experiment with written stimuli, Foucart, Martin, Moreno and Costa (2014) tested whether participants made use of the broader semantic context in Spanish to predict the gender feature of an upcoming noun. An example from the study is provided in (9). The critical word is underlined.

(9) *El pirata tenía el mapa secreto, pero nunca encontró …*
a. *el tesoro*
b. *la gruta*
… que buscaba.
'The pirate had the secret map, but he never found the[masc.] treasure/the[fem.] cave (he) was looking for.'

Like L1 speakers and simultaneous bilingual Catalan–Spanish speakers, French L1–Spanish L2 speakers showed an increased N400 on the gender-marked article if the gender did not match with the gender of the expected noun. Thus, when comprehenders encountered the article *la* (9-b), but a masculine noun was the one to be expected (9-a), this induced a larger N400 in comparison to the gender-matching article *el*. Foucart, Ruiz-Tada and Costa (2016) repeated the experiment with French L1–Spanish L2 speakers, this time with spoken stimuli and muted critical nouns. They found an enhanced long-lasting negativity after gender-mismatching articles in comparison to gender-matching articles. Moreover, in an additional lexical recognition task, expected nouns were more often falsely recognized as heard than unexpected nouns – that is, even when these nouns were not heard because they were muted. This latter finding suggests that predictions created a memory trace; for convergent findings in L1 Spanish speakers, see Foucart, Ruiz-Tada and Costa (2015). A group comparison in Foucart et al. (2014) revealed subtle differences: while the prediction effect showed up in the 300–500 ms time window for all groups, in the 500–600 ms time window it only showed up for the L1 and simultaneous bilingual group, resulting in an interaction with group. Altogether, across the written and spoken modality, L2 speakers were able to predict an upcoming noun and its gender when presented with a semantically restrictive context, such as in example (9). This was visible in form of an increased negativity on the pre-nominal article in conditions in which the gender did not match with those of a highly predictable noun in comparison to a gender-matching condition.

In an ERP experiment by Martin et al. (2013), the authors examined the prediction of a noun's phonological form in Spanish L1–English L2 speakers in comparison to an English L1 group by capitalizing on phonological agreement in English. Thus, in this experiment conditions were not favorable for prediction as Spanish has no phonological agreement. The participants in the experiment read semantically restrictive sentence contexts as in (10). The critical nouns in the two experimental conditions started with either a consonant or a vowel, so the preceding article would be *a/an*.

(10) She has a nice voice and always wanted to be …
a. a singer.
b. an artist.

When encountering the article *an* in (10-b), indicating that the more likely noun *singer* cannot follow, L1 speakers of English showed a larger N400 than for (10-a). Unlike the L1 speakers,

[2]For instance, Molinaro, Giannelli, Caffarra and Martin (2017) could show that early Basque L1–Spanish L2 speakers were more sensitive towards transparent gender cues in Spanish than early Spanish L1–Basque L2 speakers, in alignment with a persisting L1 influence.

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

the L2 speakers showed no prediction effect at the article preceding the target noun. Further L1/L2 differences were observed at noun onset. Yet, the L1 results were not fully consistent with those from a previous study by DeLong, Urbach and Kutas (2005). DeLong and colleagues reported a correlation with the noun's cloze probability – that is, the probability that this word appeared given the context as assessed via a cloze test with a separate group of participants. Instead of a graded effect, Martin et al. (2013) only found an effect of the experimental condition – that is, expected versus unexpected (for discussion, see Ito, Pickering & Corley, 2018). Furthermore, subsequent studies have questioned the replicability of pre-activation before target word onset in ERP experiments (e.g., Nieuwland, Politzer-Ahles, Heyselaar, Segaert, Darley, Kazanina, Von Grebmer Zu, Wolfsthurn, Bartolozzi, Kogan, Ito, Mézière, Barr, Rousselet, Ferguson, Busch-Moreno, Fu, Tuomainen, Kulakova, Husband, Donaldson, Kohút, Rueschemeyer, & Huettig, 2018; but cf. Nicenboim, Vasishth & Rösler, 2020). Still, if an effect of prediction, as measured by neural responses in reaction to expected and unexpected linguistic input, shows up in an L1 group but not an L2 group, this raises the question as to why prediction was more variable in the group of L2 speakers. This was also the case in the following experiment, in which participants were tested under more favorable conditions.

In a visual-world eye-tracking experiment by Ito, Pickering and Corley (2018), an English L1 and a Japanese L1–English L2 group was presented with a four-object display, showing a critical object (e.g., a cloud) and three unrelated items (an eggplant, a fork, a screw), before the target noun – underlined in (11) – was heard.

(11) The tourist expected rain when the sun went behind the cloud, but the weather got better later.

The critical object in the visual display varied depending on the experimental condition: in the target condition, the critical object corresponded to the picture of the target noun (*cloud*). In the English competitor condition, the picture of a noun that phonologically overlapped with the onset of the target noun was the critical object (*clown*). In a Japanese competitor condition, the critical object corresponded to a noun that was phonologically related (*bear*, Japanese: *kuma*) to the Japanese translation of the target noun (*cloud*, Japanese: *kumo*). Fixations on the critical object were compared to an unrelated baseline condition, in which the critical object was unrelated (*globe*). Both the L1 and L2 speakers looked more at the critical object in the target as compared to the baseline condition before noun onset, indexing successful lexical prediction. Nevertheless, the time course of anticipatory looks in the target condition was different, showing that the L2 speakers were slower (hence, this study also shows a timing difference). More importantly, only the L1 group was also more likely to look at the critical object in the English competitor condition, indicating that only the L1 group had pre-activated the nouns' phonological form. Neither the L1 nor L2 group showed any difference in looks to the critical object between the Japanese competitor and the unrelated baseline condition.

To sum up, ERP and eye-tracking experiments demonstrate that prediction at the phonological form level is particularly difficult for L2 speakers, if not impossible. In contrast, prediction at the level of semantics and syntax has been found to be achieved in L2 processing. One should note, however, that the L2 speakers in the two studies by Foucart and colleagues were familiar with gender from their other language (as were the Catalan–Spanish bilinguals in Foucart et al., 2014) and only target nouns were included that had the same gender in Spanish and French, half of them being also cognates across Spanish and French. We can only speculate that, under less favorable conditions, prediction at the level of syntax might be different or even absent.

A likely explanation as to why L2 speakers can predict an upcoming noun (e.g., Ito, Corley & Pickering, 2018) as well as a noun's gender (e.g., Foucart et al., 2014, 2016), but not its phonological form, is provided by researchers who assume that production underlies prediction, and that prediction follows the same order as production (Pickering & Gambi, 2018). The assumption that the language production system is used to predict upcoming input during comprehension has been proposed by several researchers (Dell & Chang, 2014; Huettig, 2015; Mani, Daum & Huettig, 2016; Pickering & Garrod, 2007, 2013) and is, among other things, backed up by the finding that articulatory suppression reduces prediction in L1 speakers (Martin, Branzi & Bar, 2018). Below, I briefly introduce Pickering and Gambi's account that can explain the absence of phonological form prediction in L2 processing.

Following Pickering and Gambi (2018), the same mechanisms used to produce language are also used to predict during comprehension. In their model, the stages of prediction resemble the stages of production – that is, semantics is followed by syntax and then by (phonological/orthographic) form. To predict, the language comprehender first covertly imitates the speaker's utterance and then derives the intention from this utterance. While deriving the intention, the comprehender takes into consideration non-linguistic information (e.g., shared background knowledge, visual context). Additionally, the comprehender may adjust the derived intention to compensate for differences between them and the speaker, for example, when listening to an L2 speaker (e.g., Bosker, Quené, Sanders & de Jong, 2014). The derived intention is then run through the comprehender's own production system, triggering the retrieval and build-up of production representations that constitute the comprehender's prediction of the speaker's upcoming utterance. Since each stage from conceptualization, lexical selection, morphological encoding, retrieval of phonological form up to articulation requires time and resources (Ito & Pickering, 2021), prediction might be restricted to earlier stages or remains completely absent. Notice that an earlier stage in production means a higher level in the top-down process of prediction. Ito, Gambi, Pickering, Fuellenbach and Husband (2020) could show that gender-mismatching articles in Italian elicited an earlier negativity than phonologically mismatching articles, lending support to the hypothesized order in which prediction-by-production proceeds.

To conclude, since the form level corresponds to the latest stage of prediction in accordance with prediction-by-production, differences between L1 speakers and L2 speakers should be visible especially at this level. Due to factors such as slower lexical processing and competing information, L2 speakers may lack the time to reach this level. Thus, the absence of phonological form prediction in L2 processing might be rather seen as a quantitative difference between L1 and L2 processing.

3. General discussion

At the time Kaan's review was published, in which she set out to answer what is different in L1 and L2 predictive processing, only a handful of studies existed that had tested prediction in bilingual

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

sentence processing. The current review has provided an update on what we know about L1/L2 differences in prediction after several years of extensive research. It showed that prediction studies that include a comparison between L1 and L2 processing almost always find quantitative and sometimes qualitative L1/L2 differences. As quantitative differences I considered later onsets and weaker effects of prediction in L2 processing relative to L1 processing.

Indication for a qualitative difference was provided by studies that showed a difference in the reliance of cues between L1 and L2 processing (e.g., Grüter et al., 2020; Hopp, 2015) or differences between L1 and L2 processing in the neural responses to unexpected input (Alemán Bañón & Martin, 2021). Most likely, these differences originated from a difference in the weighting of cues and/or cross-linguistic influence as a result of speakers' specific bilingual experience. L2 speakers in the study by Grüter et al. (2020) were unfamiliar with classifiers from their L1 English. While L1 speakers of Mandarin Chinese primarily relied on class membership, the L2 speakers, like the L2 Japanese speakers in the study by Mitsugi (2018), made use of the semantics of classifiers. Similarly, English L1 speakers in the study by Hopp (2015) relied on verb semantics but did not use morphological case marking in their L2 German. L1 speakers of Spanish, unlike L1 speakers of Swedish, showed a different neural response for unexpected gender marking on possessives in their L2 English than L1 speakers of English. Swedish and English mark gender on the possessor, Spanish on the possessed noun. Hence, one question that needs to be addressed when comparing prediction in L1 and L2 processing is whether a predictive cue is (similarly) encoded and similarly weighted in the L1 and L2. If not, a prediction based on that cue is likely to be absent, meaning there is no effect of prediction (van Bergen & Flecken, 2017) or, when available, another more reliable cue may be used (Grüter et al., 2020; Hopp, 2015). Similarly, if information is differently encoded on the predicted element, prediction may differ between L1 and L2 processing (Alemán Bañón & Martin, 2021). Another factor that has been highlighted as influential is the reliability of the predictive cue(s). Whether a cue is reliable or not can also depend on the experimental design and materials. Altogether, it often comes down to the language combination(s) and methodology/experimental design whether we observe an effect of prediction that is evident in L1 processing also in L2 processing.

In the next two subsections, I describe in more detail how linguistic prediction can be differently affected by individual-level factors in L2 processing. Moreover, I highlight some limitations and point out open questions and future directions.

3.1 The flow of linguistic prediction

The chart in Figure 2 exemplifies the stages of prediction for a simple sentence in which the transitive verb restricts the domain of subsequent reference to animate entities. Moreover, as shown in the literature review, the parser also takes into consideration the agent of the action. World knowledge possibly tells the parser that a girl is more likely to feed a cat than a tiger or a snake.

In the example, no cues from different linguistic domains must be integrated, nor is the sentence syntactically complex. Yet, the results from previous studies show that even in relatively simple sentences, quantitative differences between L1 and L2 processing can emerge. This can have multiple reasons. Depending on the experimental method and design, prediction of the subsequent noun can be facilitated by the visual context. If a language comprehender listens to the sentence and sees a visual scene/display that shows a cat, the concept 'cat' is likely to be pre-activated. So far, most studies that have investigated prediction in bilingual sentence processing employed the visual-world eye-tracking paradigm; out of those included in the review, only seven studies used the recording of ERPs. Hence, in most studies, non-linguistic information was available that further constrained linguistic prediction. Across methods, the prediction of 'cat' is constrained by time: the speech rate and, related to this, the time between verb onset and the onset of the noun 'cat.' In visual-world eye-tracking studies, also the preview time can influence prediction (Huettig & Guerra, 2019). Moreover, prediction can be complicated by the simultaneous presentation of competitor images. A competitor for the example *The girl feeds the cat* could be an image of another animal in addition to *cat*, such as the image of a *tiger*. This leads us to the individual-level factors influencing prediction: one such factor is competing linguistic information. Bilingual speakers likely also activate lexical representations in their other languages and experience interference from this co-activation, which costs them time, as indicated by the red hour glasses in the chart. For example, if the target language were not English but German or Spanish, the pre-nominal article would provide a gender cue that, depending on the lexical gender of the nouns depicted in the visual scene/display, could confirm a previous prediction, or rule it out. Complicating things for bilinguals, the lexical gender of nouns could be different for the other language that is co-activated. Finally, as highlighted in the chart, the target noun can share the same phonological onset in the L1 and L2, which may facilitate prediction. However, research so far indicates that L2 speakers do not reach the level of phonological form prediction (e.g., Ito, Corley & Pickering, 2018; Martin et al., 2013).

Throughout this review, two mechanisms of prediction have been mentioned: prediction-by-association and prediction-by-production. The former mechanism has been described as an automatic mechanism whose effect is rather short-lived, while prediction-by-production is more efficient but requires the availability of time and resources (Ito & Pickering, 2021; Pickering & Gambi, 2018). An increasing number of studies in recent years has provided evidence for the existence of a prediction-by-production mechanism (e.g., Ito, Pickering & Corley, 2018; Ito et al., 2020; Martin et al., 2018). More recently, Corps et al. (2022) proposed that prediction proceeds in stages with immediate associative prediction followed by more accurate prediction. Crucially, the existence of two stages was also evident in L2 processing (Corps et al., in press), so recent findings are compatible with Kaan's proposal that L1 and L2 prediction are underlyingly the same. In the chart, the two different mechanisms are represented by the two arrows.

3.2 Limitations, open questions, and future directions

The increasing number of studies in recent years has provided us with a better understanding of prediction in L2 sentence processing and new insight into the nature of L1/L2 differences. Throughout the text, I have tried to group the study outcomes into quantitative and qualitative differences. Note, however, that there is not always a clear-cut distinction – for example, L2 processing could have been just too slow for a prediction effect to emerge, which is highly dependent on the experimental design and stimuli. Is this a qualitative difference? Probably not. As regards the timing, researchers should ideally not only provide

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

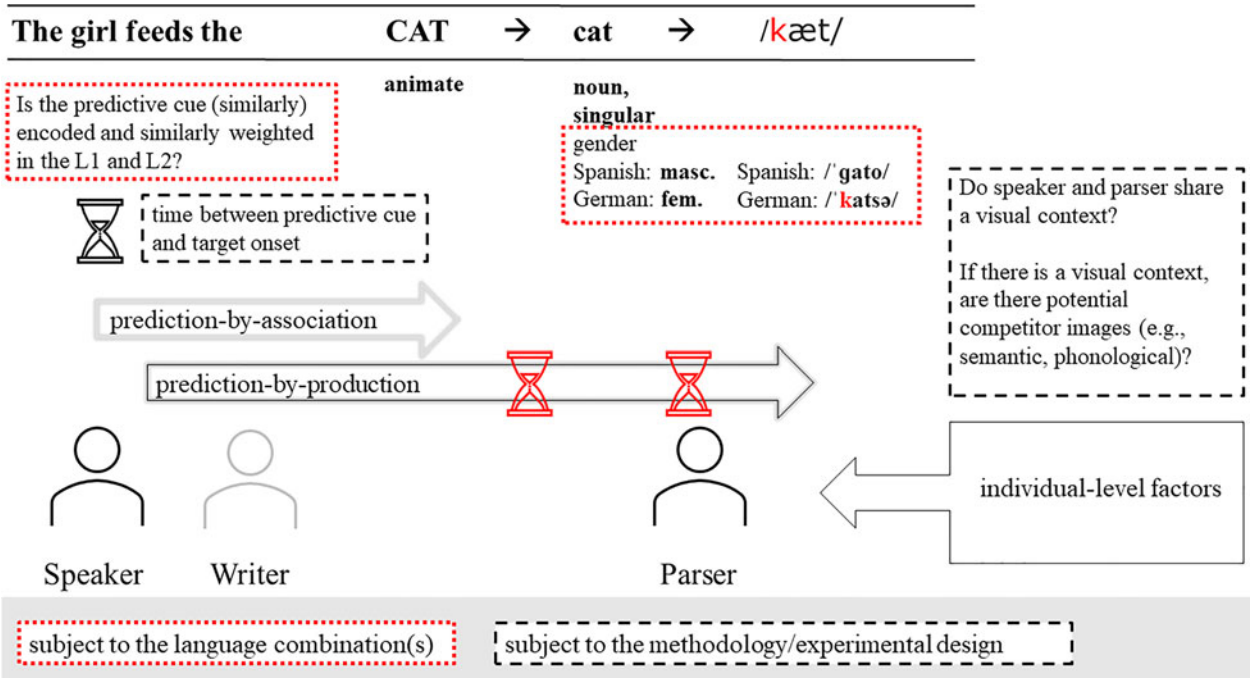


Fig. 2. The flow of prediction in sentence processing.

information about the length of the critical region for better comparability between studies, but also information about the preview time (if eye-tracking is used) and the presentation rate. As highlighted by Fernandez et al. (2020), only a few studies provide information about the presentation rate of their stimuli. While we probably do not want the stimuli in an experiment to sound artificial, we also do not want to miss an effect and, in the worst case, draw the wrong conclusion from that.

Prediction in L2 processing could also be absent because L2 speakers assign a different weighting to a cue than L1 speakers. How are we able to tell whether this is a likely explanation? As discussed by Kaan and Grüter (2021), there are several interrelated reasons why a cue can be less reliable in L2 processing. The use of a cue depends on prior experience, which can be the experience from the languages of a bilingual speaker. Moreover, also the experience within an experiment can influence the use of a cue for prediction. One option of testing for L1/L2 differences in the utility of prediction is to test bilingual speakers in both of their languages and use their L1 processing as the reference. If an effect of prediction shows up in the L1 but not the L2 and assuming that there was a fair amount of time for processing, then we have some indication that the costs of prediction did not outweigh its benefits in L2 processing. So far, only four studies include such a within-group, between-language comparison, three with adult L2 speakers (Dijkgraaf et al., 2017, 2019; Foltz, 2021a, 2021b) and one with early simultaneous bilingual children (Theimann, Kuzmina & Hansen, 2021). Such a between-language comparison has the further advantage that it offers better control of other individual-level differences affecting both languages. However, the possibilities are limited by the combination of languages and linguistic phenomena that can be directly compared. Another possibility is to compare L2 speakers with different L1 backgrounds that are otherwise matched for proficiency (and other factors) to each other. Only a few prediction studies to date include a comparison between L2 groups (Alemán Bañón & Martin, 2021; Frenck-Mestre et al., 2019; van Bergen & Flecken, 2017). Such group comparisons are particularly valuable for our understanding of the role of transfer in bilingual sentence processing. Moreover, if we are interested in the role of language exposure and the role of transfer from a developmental perspective, the investigation of prediction in bilingual children may provide "the missing piece of the puzzle" to cite the title of a recent paper by Karaca et al. (2021). So far, most studies rely on adult late bilingual speakers who had already acquired an L1 before acquiring an L2, meaning that a language system was in place. Moreover, with few exceptions, these bilingual speakers were tested in their L2, and their results compared to an L1 group that most probably knew other languages as well. To the best of my knowledge no study to date has investigated how the exposure and use of a later learned language might have had an impact on prediction in sequential bilingual speakers' L1 processing. This could be a next step forward for a more comprehensive picture of bilingual sentence processing in real-time.

Acknowledgements. I like to thank Silke Schunack, Sol Lago and Marit Westergaard as well as the reviewers for their helpful comments. The paper was completed while the author was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 101024414.

Competing interests. The author declares none.

Data Availability. An overview of studies that investigate prediction in bilingual speakers is provided on the Open Science Framework (OSF) at https://osf.io/6rjh5/. The Excel file can be filtered to search, for example, for languages and predictive cues under investigation.

References

Alemán Bañón J and Martin C (2021) The role of crosslinguistic differences in second language anticipatory processing: An event-related potentials study. *Neuropsychologia* **155**, 107797. https://doi.org/10.1016/j.neuropsychologia.2021.107797

Altmann GTM and Mirković J (2009) Incrementality and prediction in human sentence processing. *Cognitive Science* **33**, 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x

Bates E and MacWhinney B (1989) Functionalism and the competition model. In B MacWhinney and E Bates (Eds.), *The crosslinguistic study of sentence processing*. New York: Cambridge University Press, pp. 3–76.

Borovsky A, Elman JL and 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**, 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Bosch JE, Chailleux M, Yee J, Guasti MT, and Arosio F (2022) Prediction on the basis of gender and number in Mandarin-Italian bilingual children. In D Ayoun (Ed.), *Studies in Bilingualism: Vol. 63. The acquisition of gender*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 243–271.

Bosker HR, Quené H, Sanders T and de Jong NH (2014) Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language* **75**, 104–116. https://doi.org/10.1016/j.jml.2014.05.004

Chun E, Chen S, Liu S and Chan A (2021) Influence of syntactic complexity on second language prediction. In E Kaan and T Grüter (Eds.), *Bilingual Processing and Acquisition: Vol. 12. Prediction in second language processing and learning*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 70–89.

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

Clahsen H and Felser C (2006) Grammatical processing in language learners. *Applied Psycholinguistics* **27**, 3–42. https://doi.org/10.1017/S0142716406060024

Clahsen H and Felser C (2018) Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition* **40**, 693–706. https://doi:10.1017/S0272263117000250

Corps RE, Brooke C and Pickering MJ (2022) Prediction involves two stages. Evidence from visual-world eye-tracking. *Journal of Memory and Language* **122**. 104298. https://doi.org/10.1016/j.jml.2021.104298

Corps RE, Liao M, and Pickering MJ (in press). Evidence for two stages of prediction in non-native speakers: A visual-world eye-tracking study. *Bilingualism: Language and Cognition*.

Cunnings I (2017) Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition* **20**, 659–678. https://doi.org/10.1017/S1366728916000675

Dell GS and Chang F (2014) The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **369**, 20120394. https://doi.org/10.1098/rstb.2012.0394

DeLong KA, Troyer M and Kutas M (2014) Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. *Language and Linguistics Compass* **8**, 631–645. https://doi.org/10.1111/lnc3.12093

DeLong KA, Urbach TP and Kutas M (2005) Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* **8**, 1117–1121. https://doi.org/10.1038/nn1504

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

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

Duñabeitia JA, Crepaldi D, Meyer AS, New B, Pliatsikas C, Smolka E and Brysbaert M (2018): MultiPic: A standardized set of 750 drawings with norms for six European languages, *Quarterly Journal of Experimental Psychology* **71**, 808–816. http://dx.doi.org/10.1080/17470218.2017.1310261

Dussias PE, Valdés Kroff JR, Guzzardo Tamargo RE and Gerfen C (2013) When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition* **35**, 353–387. https://doi.org/10.1017/S0272263112000915

Fernandez LB, Engelhardt PE, Patarroyo AG and Allen SE (2020) Effects of speech rate on anticipatory eye movements in the visual world paradigm: Evidence from aging, native, and non-native language processing. *Quarterly Journal of Experimental Psychology* **73**, 2348–2361. https://doi.org/10.1177/1747021820948019

Foltz A (2021a) Adaptation in predictive prosodic processing in bilinguals. *Frontiers in Psychology* **12**:661236. https://doi.org/10.3389/fpsyg.2021.661236

Foltz A (2021b) Using prosody to predict upcoming referents in the L1 and the L2: The role of recent exposure. *Studies in Second Language Acquisition* **43**, 753–780. https://doi.org/10.1017/S0272263120000509.

Foucart A (2021) Language prediction in second language: Does language similarity matter? In E Kaan and T Grüter (Eds.), *Bilingual Processing and Acquisition: Vol. 12. Prediction in second language processing and learning*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 92–114.

Foucart A, Martin CD, Moreno EM and Costa A (2014) Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition* **40**, 1461–1469. https://doi.org/10.1037/a0036756

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

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

Frenck-Mestre C, Kim SK, Choo H, Ghio A, Herschensohn J and Koh S (2019) Look and listen! The online processing of Korean case by native and non-native speakers. *Language, Cognition and Neuroscience* **34**, 385–404. https://doi.org/10.1080/23273798.2018.1549332

Grüter T, Lau E and 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**, 221–234. https://doi.org/10.1080/23273798.2019.1648840

Grüter T, Lew-Williams C and Fernald A (2012) Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research* **28**, 191–215. https://doi.org/10.1177/0267658312437990

Grüter T and Rohde H (2021) Limits on expectation-based processing: Use of grammatical aspect for co-reference in L2. *Applied Psycholinguistics* **42**, 51–57. https://doi.org/10.1017/S0142716420000582

Grüter T, Takeda A, Rohde H and Schafer AJ (2018) Intersentential coreference expectations reflect mental models of events. *Cognition* **177**, 172–176. https://doi.org/10.1016/j.cognition.2018.04.015

Henry N, Hopp H and Jackson CN (2017) Cue additivity and adaptivity in predictive processing. *Language, Cognition and Neuroscience* **32**, 1229–1249. https://doi.org/10.1080/23273798.2017.1327080

Henry N, Jackson CN and Hopp H (2022) Cue coalitions and additivity in predictive processing: The interaction between case and prosody in L2 German. *Second Language Research* **38**, 397–422. https://doi.org/10.1177/0267658320963151.

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

Hopp H (2021) Prediction and grammatical learning in second language sentence processing. In E Kaan and T Grüter (Eds.), *Bilingual Processing and Acquisition: Vol. 12. Prediction in second language processing*

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

and learning. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 168–185.

Hopp H and Lemmerth N (2018) Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition* **40**, 171–199. https://doi.org/10.1017/S0272263116000437

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, Audring J and Jackendoff R (2022) A parallel architecture perspective on pre-activation and prediction in language processing. *Cognition* **224**:105050. https://doi.org/10.1016/j.cognition.2022.105050

Huettig F and 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 and Pickering MJ (2019) Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Sciences* **23**, 464–475. https://doi.org/10.1016/j.tics.2019.03.008

Ito A, Corley M and Pickering MJ (2018) A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition* **21**, 251–264. https://doi.org/10.1017/S1366728917000050

Ito A, Gambi C, Pickering MJ, Fuellenbach K and Husband EM (2020) Prediction of phonological and gender information: An event-related potential study in Italian. *Neuropsychologia* **136**, 107291. https://doi.org/10.1016/j.neuropsychologia.2019.107291

Ito A and Pickering MJ (2021) Automaticity and prediction in non-native language comprehension. In E Kaan & T Grüter (Eds.), *Bilingual Processing and Acquisition: Vol. 12. Prediction in second language processing and learning*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 26–46.

Ito A, Pickering MJ and 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

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

Kaan E and Grüter T (2021) Prediction in second language processing and learning: Advances and directions. In E Kaan and T Grüter (Eds.), *Bilingual Processing and Acquisition: Vol. 12. Prediction in second language processing and learning*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 2–24.

Kaan E, Kirkham J and Wijnen F (2016) Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition* **19**, 1–18. https://doi.org/10.1017/S1366728914000844

Karaca F, Brouwer S, Unsworth S and Huettig F (2021) Prediction in bilingual children: The missing piece of the puzzle. In E Kaan & T Grüter (Eds.), *Bilingual Processing and Acquisition: Vol. 12. Prediction in second language processing and learning*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 116–137.

Kim H and Grüter T (2021) Predictive processing of implicit causality in a second language. *Studies in Second Language Acquisition* **43**, 133–154. https://doi.org/10.1017/S0272263120000443

Koch EM, Bulté B, Housen A and Godfroid A (2021) Using verb morphology to predict subject number in L1 and L2 sentence processing: A visual-world eye-tracking experiment. *Journal of the European Second Language Association* **5**, 115–132. https://doi.org/10.22599/jesla.79

Kuperberg GR and Jaeger TF (2016) What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* **31**, 32–59. https://doi.org/10.1080/23273798.2015.1102299

Lago S, Mosca M and Stutter Garcia A (2021) The role of crosslinguistic influence in multilingual processing: Lexicon versus syntax. *Language Learning* **71**, 163–192. https://doi.org/10.1111/lang.12412

Lemmerth N and Hopp H (2019) Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Language Acquisition* **26**, 21–45. https://doi.org/10.1080/10489223.2017.1391815

Lew-Williams C and Fernald A (2010) Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language* **63**, 447–464. https://doi.org/10.1016/j.jml.2010.07.003

MacWhinney B (2008) A unified model. In P Robinson and NC Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Press, pp. 343–371.

Mani N, Daum MM and Huettig F (2016) "Proactive" in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. *Quarterly Journal of Experimental Psychology* **69**, 2189–2201. https://doi.org/10.1080/17470218.2015.1111395

Mani N and Huettig F (2014) Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology* **126**, 264–279. https://doi.org/10.1016/j.jecp.2014.05.004

Martin CD, Branzi FM and Bar M (2018) Prediction is production: The missing link between language production and comprehension. *Scientific Reports* **8**, 1079. https://doi.org/10.1038/s41598-018-19499-4

Martin CD, Thierry G, Kuipers J-R, Boutonnet B, Foucart A and Costa A (2013) Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* **69**, 574–588. https://doi.org/10.1016/j.jml.2013.08.001

Meir N, Parshina O and Sekerina IA (2020). The interaction of morphological cues in bilingual sentence processing: An eye-tracking study. In MM Brown and A Kohut (Eds.), *Proceedings of the 44th Annual Boston University Conference on Language Development* (pp. 376–389). Somerville MA: Cascadilla Press.

Mishra RK, Singh N, Pandey A and Huettig F (2012) Spoken language-mediated anticipatory eye-movements are modulated by reading ability – Evidence from Indian low and high literates. *Journal of Eye Movement Research* **5**, 1–10. https://doi.org/10.16910/jemr.5.1.3

Mitsugi S (2017) Incremental comprehension of Japanese passives: Evidence from the visual-world paradigm. *Applied Psycholinguistics* **38**, 953–983. https://doi.org/10.1017/S0142716416000515

Mitsugi S (2018) Generating predictions based on semantic categories in a second language: A case of numeral classifiers in Japanese. *International Review of Applied Linguistics in Language Teaching* **58**, 323–349. https://doi.org/10.1515/iral-2017-0118

Mitsugi S (2022) Polarity adverbs facilitate predictive processing in L2 Japanese. *Second Language Research* **38**, 869–892. https://doi.org/10.1177/02676583211000837.

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

Molinaro N, Giannelli F, Caffarra S and Martin C (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition* **164**, 61–73. https://doi.org/10.1016/j.cognition.2017.03.012

Nicenboim B, Vasishth S and 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 MS, 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 DJ, Rousselet GA, Ferguson HJ, Busch-Moreno S, Fu X, Tuomainen J, Kulakova E, Husband EM, Donaldson DI, Kohút Z, Rueschemeyer SA, and 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

Perdomo M and Kaan E (2021) Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research* **37**, 349–375. https://doi.org/10.1177/0267658319879196

Peters R, Grüter T, and Borovsky A (2018) Vocabulary size and native speaker self-identification influence flexibility in linguistic prediction among adult bilinguals. *Applied Psycholinguistics* **39**. 1439–1469.

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press

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

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

Pickering MJ and Garrod S (2013) An integrated theory of language production and comprehension. *Behavioral and Brain Sciences* **36**, 329–347. https://doi.org/10.1017/S0140525X12001495

Radošević T, Malaia EA, and Milković M (2022) Predictive processing in sign languages: A systematic review. *Frontiers in Psychology* **13**:805792. https://doi.org/10.3389/fpsyg.2022.805792

Ryskin R, Levy RP and Fedorenko E (2020). Do domain-general executive resources play a role in linguistic prediction? Re-evaluation of the evidence and a path forward. *Neuropsychologia* **136**, 107258. https://doi.org/10.1016/j.neuropsychologia.2019.107258

Schlenter J and Felser C (2021) Second language prediction ability across linguistic domains: Evidence from German. In E Kaan and T Grüter (Eds.), *Bilingual Processing and Acquisition: Vol. 12. Prediction in second language processing and learning*. Amsterdam, Philadelphia: John Benjamins Publishing Company, pp. 48–68.

Theimann A, Kuzmina E, and Hansen P (2021) Verb-mediated prediction in bilingual toddlers. *Frontiers in Psychology* **12**:719447. https://doi.org/10.3389/fpsyg.2021.719447

Van Bergen G and 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

https://doi.org/10.1017/S1366728922000736 Published online by Cambridge University Press