BASIC SCIENCE ARTICLE

Predictive sentence processing in L2 and L1

What is different?*

Edith Kaan
University of Florida

There is ample evidence that native speakers anticipate upcoming information at various levels during sentence comprehension. In contrast, some studies on late second-language (L2) learners support the view that L2 learners do not anticipate information during processing, or at least, not to the same extent as native speakers do. In the current paper, I propose that native and L2 speakers are underlyingly the same as far as sentence processing mechanisms are concerned, and that potential differences in anticipatory behavior can be accounted for by the same factors that drive individual differences in native speakers; in particular, differences in frequency biases, competing information, the accuracy and consistency of the lexical representation, and task-induced effects. Suggestions for future research are provided.

Keywords: predictions, sentence processing, second-language

1. Predictive processing in native and non-native sentence processing[1]

The field of sentence processing, and cognitive science in general (Bar, 2009; Clark, 2013), has seen a recent surge in interest in predictive processing. In the past few years, quite a few papers and presentations have proposed that the language comprehender is not only integrating each word into the preceding structure as it comes in but also forming predictions as to which syntactic structure, word category, or lexical item comes next (e.g., Altmann & Mirkovic, 2009; Jaeger & Snider, 2013; Kutas, DeLong & Smith, 2011; Levy, 2008; MacDonald, 2013; Pickering & Garrod, 2011, 2013; Van Petten & Luka, 2012). When the input does not match the predictions, this mismatch is used to adjust future predictions and minimize the chance of future errors (Jaeger & Snider, 2013). This continuous adjustment of what to expect can account for language learning (Chang, Dell, & Bock, 2006) and for speaker alignment and turn taking

Linguistic Approaches to Bilingualism 4:2 (2014), **257–282**. DOI 10.1075/lab.4.2.05kaa
ISSN 1879–9264 / E-ISSN 1879–9272

(Pickering & Garrod, 2011, 2013), hence optimizing communication (Jaeger & Snider, 2013).

Although the idea of predictive, top-down processing is not new (e.g., Abney & Johnson, 1991; Crain & Fodor, 1985; Gibson, 1998) and many studies assume a predictive component (e.g., Fischler & Bloom, 1979), new computational approaches such as Bayesian statistics, and experimental paradigms such as visual world eye-tracking paradigms and electrophysiology, have enabled investigators to directly tap into anticipatory processing. Anticipatory behavior can be attested when participants move their eyes to an object before it is explicitly referred to in the auditory input. For instance, Altmann and Kamide (1999) had participants view a scene with a boy, a toy car, a toy train, and a cake, and recorded eye movements while participants listened to sentences such as *The boy will eat the cake*. After hearing *eat*, the listeners moved their eyes to the edible object in the scene (cake) even before the actual word *cake* was presented in the input. This suggests that participants use their knowledge of the verb *to eat* to anticipate what will come next in the input. In research using Event-Related brain Potentials (ERPs), indices of incongruency can be observed at determiners or gender-markings before participants read or hear the actual lexical item that is unexpected given the preceding context (DeLong, Urbach & Kutas, 2005; Van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; Wicha, Moreno & Kutas, 2003, 2004). In addition, electroencephalography (EEG) signals in specific frequency bands are modulated by the degree to which a particular word can be predicted, before this word is presented in the input (Dikker & Pylkkänen, 2013; Molinaro, Barraza & Carreiras, 2013).

A growing number of eye-tracking and electrophysiological studies have shown that native speakers of a language can predict upcoming information in detail and at various levels of representation including the grammatical gender of a word (Van Berkum et al., 2005), the phonological form of the word (DeLong et al., 2005), the visual word form (Dikker, Rabagliati, Farmer, & Pylkkänen, 2010; Dikker, Rabagliati, & Pylkkänen, 2009;), the syntactic category (Kim & Gilley, 2013; Lau, Stroud, Plesch, & Phillips, 2006), the arguments of the verb (Boland, 2005), or an object gap position in *wh* dependencies (Dickey, Choy, & Thompson, 2007; Dickey & Thompson, 2009; Sussman & Sedivy, 2003). These predictions can be based on real-world knowledge (Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003a; Van Berkum et al., 2005; Wicha et al., 2003, 2004), syntactic information such as gender, case, or preceding syntactic structure (Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Dickey et al., 2007; Dickey & Thompson, 2009; Kamide, Scheepers, & Altmann, 2003b; Lau et al., 2006; Lew-Williams & Fernald, 2010; Sussman & Sedivy, 2003), properties of the verb (Boland, 2005; Kamide et al., 2003a;), or prosody (Nakamura, Arai, & Mazuka, 2012; Weber, Grice, & Crocker, 2006).

Whereas most native speakers appear to use information rapidly to anticipate upcoming information, studies on second-language (L2) learners suggest that non-native speakers do not anticipate to the same extent as native speakers, even though they know the specific rules and words used when probed off-line (Dussias, Valdés Kroff, Guzzardo Tamargo & Gerfen, 2013; Grüter, Lew-Williams & Fernald, 2012; Grüter & Rohde, 2013; Hopp, 2013; Lew-Williams & Fernald, 2010; Martin et al., 2013). For instance, some studies using a visual world eye-tracking paradigm suggest that moderately proficient English L2 learners of Spanish do not use gender information to predict the upcoming noun, in contrast to native monolingual speakers of Spanish (Lew-Williams & Fernald, 2010; Dussias et al., 2013). Spanish has two syntactic genders (feminine and masculine), which are marked at the article preceding the noun. In studies using a visual world paradigm, participants see a display with pictures. Two of the pictures depict the meaning of nouns that either have the same syntactic gender, or a different syntactic gender. Participants listen to sentences or instructions that contain a gender-marked determiner (*Encuentra el/la…*, 'Find the$_{\text{MASC}}$/ the$_{\text{FEM}}$…') while their eye movements are recorded. Lew-Williams and Fernald (2010) reported that when the display contained pictures depicting nouns that were different in gender, native Spanish speakers moved their eyes to the target object when they heard the disambiguating gender marked determiner; that is, before they heard the noun. Adult English learners of Spanish, however, did not use the determiner as a predictive cue, even though they knew the rules of gender concord in Spanish and the correct gender of the nouns that were used (Grüter et al., 2012). This suggests that although L2 speakers have the relevant L2 knowledge, they cannot use this knowledge to anticipate upcoming information. Quite a number of studies support the view that L2 speakers do not anticipate, at least not to the extent that native speakers do (Dussias et al., 2013; Grüter et al., 2012; Grüter & Rohde, 2013; Hopp, 2013; Martin et al., 2013), although this may depend on L2 proficiency (Dussias et al., 2013; Hopp, 2013), the L2 speakers' native language, and the kind of information serving as a potential predictive cue (Dussias et al., 2013).

The observation that L2 learners differ from native speakers with respect to anticipatory processing is a novel and rather consistent finding, with potentially wide ramifications. For instance, Grüter and Rohde (2013) propose the RAGE model (Reduced Ability to Generate Expectations during language processing) for L2 learners, which accounts for differences and similarities observed between L2 and native processing at the sentence and discourse level. An earlier paper, Kaan, Dallas, & Wijnen (2010), argued that an L2 speaker's inability to actively predict upcoming information can account for some previously observed differences between L2 and native syntactic processing, in particular the absence of early left anterior negativity (ELAN) effects (Hahne, 2001; Hahne & Friederici,

2001), delayed filled-gap effects (Dallas, 2008), and the absence of intermediate trace effects in (non-immersed) intermediate-level L2 learners (Marinis, Roberts, Felser, & Clahsen, 2005). However, some studies do observe native-like predictive processing in L2 learners (Dussias et al., 2013; Hopp, 2013). Therefore, claiming that L2 and native speakers differ in sentence processing in that native speakers actively predict and L2 speakers do not runs counter to the available evidence; neither does it account for why L2 learners anticipate information before it appears in the input in some cases but not in others. An example comes from a study by Dussias et al. (2013) in which Italian learners of Spanish did not anticipate a noun when they heard a masculine determiner, but they did when they heard a feminine determiner. I propose that native speakers and L2 learners do not differ in the nature of the predictive mechanisms or in the way these mechanisms are employed, but in what drives these mechanisms. For practical reasons, I will focus on sentence comprehension although I acknowledge that the comprehension and production processes are closely tied (MacDonald, 2013; Pickering & Garrod, 2013). I will discuss factors that affect predictive processing in native speaker sentence processing, showing how each of these factors may be different in non-native speakers and how this can affect their anticipatory processing. Differences in predictive processing between native and non-native speakers can therefore be accounted for by the same factors that drive individual differences in native speakers, without assuming that L2 speakers are inherently different from native speakers in terms of predictive processing. I will conclude with a comparison of other proposals on L2 processing and make suggestions for future research.

2. Factors modulating prediction

A variety of current models of (monolingual) native sentence processing incorporate a predictive component (e.g., Gibson, 1998, 2000; Hale, 2006; Levy, 2008; MacDonald, 2013). I will not try to distinguish among these models; instead, I will assume a rather general approach, mainly based on ideas put forward by researchers working in a connectionist tradition (e.g., Elman, 1991; MacDonald, 2013; MacDonald, Pearlmuttter, & Seidenberg, 1994; Tabor & Tanenhaus, 1999). Based on a lifetime of exposure to their native language, native speakers of a language have learned to associate certain words, word categories, syntactic frames and rules, building a database of relative frequencies of occurrence and associations. These (co-occurrence) frequencies are not static but instead change over time when interacting with other users of the language. While listening or reading, language users construct a syntactic, semantic, and phonological/orthographical representation of the input, assigning likelihoods to the various ways in which

the current utterance can be continued. If a particular continuation is strongly predicted (i.e., is very likely) but is contradicted by the input or not found in it, processing difficulty ensues.[2] In this situation, the prediction error, that is, the discrepancy between the predicted and actual input, is used to adjust the options available for upcoming material and to update in memory the frequencies of the structures and associations in question. This enables language users to learn and to adapt to their interlocutor.

Assuming this framework, the degree to which a listener or reader can predict upcoming words and structures is a function of several interdependent sets of factors, all of which are subject to individual differences.[3] The factors I will discuss below are: (1) the frequency information stored; (2) competing information; (3) the accuracy and consistency of the lexical information retrieved; (4) task-induced processes and strategies; and (5) other factors such as motivation, resources, and cognitive control. Each of these factors has been previously shown to affect language comprehension, and thus it is not unique to predictive processing. It is therefore not necessary to invoke special mechanisms to account for predictive processing and the difference therein between individuals and groups. These factors are relevant for prediction because they modulate the degree to which information can be pre-activated in anticipation of the input as well as the degree to which prediction errors can be used to adjust future prediction. These factors are interrelated and they interact, e.g., better cognitive control could compensate for having to deal with more competing information. Therefore, it is unlikely that variation in a single factor underlies the differences observed between populations as well as the differences within populations. Individual variation in any of the above factors can lead to differences in predictive processing within native speakers, within second-language learners, as well as between native speakers and second-language learners. I will discuss each of these factors below and show how some existing L2 data can be accounted for without the need to assume that L2 speakers are inherently different from native speakers in the predictive mechanisms they use.

2.1 Frequency information

Language users are continually exposed to language and, on the basis of what they encounter, they store information about the likelihood of a certain word or construction occuring in the context of another. These frequencies are used to predict. A rather trivial example is that the verb *eat* is more likely to occur with an edible object than with an inedible object. This can account for Altmann and Kamide's (1999) findings that listeners anticipatorily move their eyes to a picture of a cake when they hear *The boy will eat the…*. Another example of the effect of frequency information is the subcategorization bias of verbs. Scholars working

in the constraint-based lexicalist framework have provided ample evidence that the relative frequencies of subcategorization frames and argument types of a verb affect processing preferences (e.g., Boland, Tanenhaus, Garnsey, & Carlson, 1995; MacDonald et al., 1994; Stowe, Tanenhaus, & Carlson, 1991; Wilson & Garnsey, 2009). This information is also used in a predictive manner, as shown by Boland (2005). Boland had participants listen to sentences such as *One window was broken, so the handyman mentioned/fixed it…* . When the verb was a dative verb, such as *mentioned*, participants moved their eyes to the picture of the recipient (*owners*) before the word *owners* appeared in the input. This was not the case for verbs such as *fixed*, which do not take a recipient argument.

Therefore, frequency information is an important source of predictive processing in native speakers. Not surprisingly, non-native speakers differ from native speakers in the nature of the language they are exposed to and the amount of exposure they receive. The frequency biases that a language user stores and uses are very much determined by the language user's linguistic environment. For instance, subcategorization preferences for particular verbs may differ among varieties of English (Schilk, 2011). Therefore, these preferences may not be the same for the L2 English learners who are tested in a study as they are for the native English speakers they are compared with. Additionally, it is often the case that non-immersed L2 speakers exclusively hear the language as produced by other L2 speakers. This language is likely to be affected by properties of the learners' native language and/or by more general properties of an interlanguage (Selinker, 1972). Moreover, the exposure an L2 learner in a non-immersion situation receives is mostly confined to textbooks and other formal language, clearly pronounced or written, which is different from the input that a native language learner receives. As a result, L2 speakers may differ from L1 speakers in terms of how frequently learners are exposed to particular combinations of words and constructions, leading to different frequency biases, and hence, differences in predictive processing.

Evidence that L2 speakers differ from native speakers in terms of frequency biases comes from a study on subcategorization biases. Dussias, Marful, Gerfen, and Bajo (2010) asked advanced, late Spanish learners of English to complete sentences such as *John admitted…*, to determine whether the L2 learners preferred the verb to take a direct object noun phrase, or a sentential complement. Results were compared to completion data obtained from native English speakers (Garnsey, Pearlmutter, Myers, & Lotocky, 1997). Dussias et al. found that the Spanish learners of English assigned a different bias than the native English speakers to 39 of the 100 verbs tested. In 10 of these 39 cases, the learners transferred the bias of the corresponding verb in Spanish to the English verb. In 22 cases, the learners assigned a bias that was different from both the bias they assigned to the Spanish translation and the bias that native English speakers assigned. These data suggest

that the relative subcategorization biases in L2 learners do not completely correspond to those of native speakers.

The difference in the frequency information stored may have consequences for L2 sentence processing and for the kind of predictions L2 learners make: a different frequency representation may either lead to weaker predictions, or to predictions that are different from those native speakers make. Unfortunately, no study has investigated the effect of subcategorization information on anticipatory looks during listening in L2 speakers, or on ERP components before the point of disambiguation. Nevertheless, results from reading studies support the view that L2 learners are similar to native speakers in that they use frequency information to predict upcoming information, and that apparent differences in processing can be attributed to differences in frequency information. For example, if a verb has a strong bias for a sentential complement, as is the case for the verb *admitted* (Wilson & Garnsey, 2009), native speakers take longer to read the words that make clear that the complement is not a clause (e.g., at *because she* in *The ticket agent admitted the mistake because she…*); if a verb has a bias for a transitive frame with a direct object (such as the verb *confirmed* in *The CIA confirmed the rumor could mean…*,) encountering the finite verb downstream (*could*) leads to an increase in reading times (Wilson & Garnsey, 2009). The increase in reading times at the point of disambiguation for non-preferred versus preferred continuations is also known as a *garden path effect*. Dussias and Cramer Scaltz (2008) found differences between Spanish L2 speakers of English and native English speakers in the size of the garden-path effect when the verb had a bias for a sentential complement (e.g. *admitted*); whereas the native English speakers showed longer reading times for the non-preferred direct object continuation (*because she*) than the sentential complement continuation (*could mean*), the L2 learners showed no differences between both types of continuation. Dussias and Cramer Scaltz showed that this difference between the groups could be attributed to differences in knowledge of subcategorization biases. When L2 data were restricted to trials for which the L2 speakers indicated the same biases as native speakers, the L2 speakers showed the same pattern as the native speakers.

The study mentioned above is an example of a study in which L2 speakers showed a smaller garden path effect than native speakers. If L2 speakers are like native speakers in that they make predictions concerning upcoming structures, but different from native speakers in the frequency biases stored, one also expects situations in which L2 speakers show stronger garden path effects, or garden paths effects in a different direction, when compared with to native speakers. One example comes from a self-paced reading study by Lee, Lu, and Garnsey (2013). Lee et al. investigated the effect of verb bias on the resolution of direct object/sentential complement ambiguities in English. The condition that is of most interest here

is the condition in which the main clause verb was biased towards a sentential complement. The sentential complement was either introduced by the complementizer *that* or not, e.g., *The ticket agent admitted (that) the mistake might be hard to correct.* The native English speakers in the study read the disambiguating auxiliary (*might*) equally fast, regardless of whether *that* was present or not. In English, verbs that are strongly biased towards a sentential complement often occur without *that*; hence, the presence of *that* was no additional cue to the native English speakers that the continuation was a sentential complement. Low-proficient Korean L2 learners of English, on the other hand, read the auxiliary (*might*) faster when *that* was present than when *that* was absent. Also, when *that* was present, the low-proficient L2 learners read the auxiliary faster when the verb bias was toward a sentential complement than when the verb bias was towards a direct object continuation. This suggests that the low-proficient L2 learners were familiar with English verb biases and used verb bias information during reading, but had not yet incorporated the frequency cues associated with the use of the complementizer *that*. This resulted in a processing pattern that was different from that of native speakers. Advanced Korean speakers of English performed similarly to native English speakers, suggesting that they had incorporated the frequency information and they could combine the predictive cues.

The Dussias and Cramer Scaltz (2008) data and the Lee et al. (2013) findings would be hard to account for by assuming that L2 speakers do not predict whereas native speakers do. Instead, the difference between L2 and native speakers lies in the frequencies of use that underlie predictive behavior. Differences in frequency biases between the native and second-language speakers of a language may, however, be a rather small factor in explaining differences in predictive processing between native speakers and L2 learners. First, some studies show that L2 speakers are able to acquire frequency biases of verbs similar to native speakers', even in non-immersion situations (Dussias et al., 2010; Martinez-Garcia & Wulff, 2012; Wulff, Lester, & Martinez-Garcia, in press). Second, there is some evidence that L2 speakers do not carry over subcategorization biases from their L1 (Flett, Branigan, & Pickering, 2013; Gries & Wulff, 2005); however, this does not mean that differences in frequency biases between a speaker's L2 and L1 do not affect L2 processing. The L1 biases may compete with the L2 biases during processing. Additionally, subcategorization preferences and other information may not always be retrieved accurately and consistently during L2 processing. I will discuss these two issues below.

2.2 Competing information

Not only are L2 speakers exposed to different input than native speakers of the language are, L2 learners are also different from many of the native speakers with whom they are compared in that L2 speakers need to deal with a dominant, native language in addition to their second language. Many studies on lexical processing have shown that early and late bilinguals activate lexical information in both languages, even in contexts in which only one language is relevant. This means that when bilinguals process words in one language, they also activate related words in their other language, especially if this other language is dominant (e.g., Dijkstra, Grainger, & van Heuven, 1999; Spivey & Marian, 1999; Van Hell & Dijkstra, 2002). This co-activation is even enhanced in the auditory domain: due to misperception and misrepresentation of phonological categories in the L2, more words in both L1 and L2 are perceived as similar to the input and become activated (e.g., Weber & Cutler, 2004). This massive activation of competing information may impair predictive processing.

Prediction becomes hard if multiple, almost equally probable alternative continuations are possible. For instance, if an English speaker hears the fragment *Pick up the can…* in a context in which there is both a candy and a candle, the target object cannot be predicted until more information is available. Averaging across trials, participants will fixate on the target object much later in the situation in which a competitor is present compared to situations in which only one object (candle or candy) is available (Spivey-Knowlton, Tanenhaus, Eberhard, & Sedivy, 1998). The same situation holds for visual world eye-tracking experiments on syntactic gender. If the display contains pictures that correspond to nouns that have the same syntactic gender (Dahan et al., 2000; Dussias et al., 2013; Hopp, 2013; Lew-Williams & Fernald, 2010), participants may either not actively predict, or maintain multiple, competing predictions in parallel, which does not lead to an obvious preference for the intended target object before it appears in the input. Due to the co-activation of information in both languages, it is therefore likely that language learners suffer more from competition and they are less likely to make specific predictions, or are slower to do so.

Some evidence for interlingual competition during L2 syntactic processing comes from a study by Frenck-Mestre and Pynte (1997) using eye-tracking during reading. Frenck-Mestre and Pynte report that French L2 learners of English had increased reading times for verbs like *obeyed* versus *barked* in *Every time the dog obeyed/barked the pretty little girl showed…..* The verb *obeyed* is optionally transitive in English but intransitive in French; *barked* is intransitive in both languages. The increase in reading time for the L2 speakers at *obeyed* suggest that the L2 learners experienced interference from competition of the subcategorization

biases associated with the equivalent of *obey* in their native language. The L2 speakers, however, did not differ from the native speakers regarding an increase in reading times at the disambiguating word downstream (*showed* in the example above). This suggests that competition was overcome quickly enough and did not differentially affect the (predictive) processing of the upcoming structure (see also Jacob, 2009, for short-lived competition from L1 word order).

Suggestive evidence that interlingual competition can affect the use of information as a predictive cue in L2 comes from a study by Dussias, Valdés Kroff, Guzzardo, and Gerfen (2013). This study used a visual world eye-tracking paradigm to test the use of gender information in Italian learners of Spanish. Italian learners of Spanish of intermediate proficiency launched anticipatory eye movements to the feminine object in the display after hearing the Spanish feminine article *la*, but they did not show anticipatory eye movements when hearing the masculine determiner *el*. Although the experiment was not designed to test interlingual competition, one of the explanations that the authors provide for the observed pattern is that Spanish *el* corresponds to two determiners in Italian, *il* and *lo*; whereas Italian only has one feminine determiner, mapping onto Spanish *la*. The activation of both Italian forms when hearing the Spanish masculine determiner may therefore have impeded the predictive use of the Spanish determiner *el* in these learners.

Bilinguals and language learners may activate more information during processing than monolingual native speakers, which may impede or slow down the formation of predictions. This can potentially account for the delayed effects observed in some L2 studies on *wh*-movement (e.g., Dallas, 2008; Dekydtspotter, Schwartz, & Sprouse, 2006; Felser, Cunnings, Batterham, & Clahsen, 2012), and the absence of predictive effects in (lower-proficiency) L2 learners in paradigms in which the time delay between the predictive element (e.g., a gender marked determiner) and the target word is relatively short (e.g., Grüter et al., 2012; Lew-Williams & Fernald, 2010). The evidence to date suggests that interlingual competition is rather short-lived (Frenck-Mestre & Pynte, 1997; Jacob, 2009). More research is needed to see to what extent interlingual competition affects predictive behavior, and what the time-course of this effect is, e.g., by systematically varying the delay between the predictive information and the target.

2.3 Accuracy and consistency of lexical information

Even among monolingual speakers, individuals differ in their predictive use of language information. First, adult native speakers anticipate to a lesser extent when their vocabulary is smaller (Borovsky, Elman, & Fernald, 2012). In addition, low-literates do not anticipate upcoming information in spoken language as much as high-literates (Mishra, Singh, Pandey, & Huettig, 2012). This suggests

that language experience and practice in reading and writing contributes to the ability to anticipate. Furthermore, native German speakers who were faster in accessing and using lexical items were also faster at using gender-marked determiners in an anticipatory fashion (Hopp, 2013).

An underlying source of these effects may be what Perfetti calls the quality of the lexical representation. The quality of a lexical representation refers to the stability and accuracy of the language user's knowledge of a word's form, meaning, and use (Perfetti, 2007; Perfetti & Hart, 2001). For instance, having a low-quality lexical representation for *gait* and *gate* can mean that the spelling between the two words is often confused. A high-quality representation implies that the information is accurate and consistent, that potential confusion with other lexical items is short-lived, and hence, that the lexical information can be retrieved easily (Perfetti & Hart, 2001). Although, to my knowledge, no study has looked directly at lexical quality in native speakers in relation to predictive processing, the data referred to above suggest a strong relation. More extensive reading and writing experience contributes to more high-quality word representations (Perfetti & Hart, 2001), which in turn can be used to predict upcoming information. Having a stable representation of a word's spelling, pronunciation, meaning, and grammatical characteristics (e.g., gender, inflection, number and type of arguments, syntactic position), and strong connections between these aspects, will help to retrieve this information consistently and easily (Perfetti & Hart, 2001), which is a prerequisite for using this information in a predictive manner.[4]

As discussed in Section 2.1, L2 speakers have typically had less exposure to their second language than native speakers of that language have had. Less exposure leads to an underrepresentation of low-frequency words and weaker associative connections (Gollan et al., 2011; Gollan, Montoya, Cera, & Sandoval, 2008). In addition, as we have seen above, bilinguals activate both their languages, which implies that lexical items are harder to distinguish from each other in a bilingual than in a monolingual lexicon. It is therefore likely that L2 learners have more lower-quality lexical representations than native monolingual speakers do, leading to a less consistent and more effortful retrieval of the information. In support of this, a large number of studies have shown that bilinguals are slower at picture naming and experience more tip-of-the-tongue effects (Gollan & Acenas, 2004; Gollan, Montoya, Fennema-Notestine & Morris, 2005). In addition, non-native speakers are less able to distinguish frequent collocative expressions such as *social services* from atypical expressions such as *permanent horror* (Siyanova & Schmitt, 2008), and are less sensitive to the order of fixed expressions such as *bread and butter* than native speakers (Siyanova & Conklin, 2008). This suggests that even though non-native speakers may have similar frequency biases to native speakers, their representation of the information is less consistent than native speakers'.

Thus far, only one study has explicitly tested the relation between consistency of information stored for a lexical item and the ability to use this information predictively during sentence processing (Hopp, 2013). Hopp (2013) tested English learners of German in a visual world eye-tracking paradigm on gender processing. The learners of German were divided into two groups: those who systematically produced the same gender for a particular noun, and those who showed inconsistencies. The learners who were consistent in their assignment of gender used gender information in a predictive manner and were indistinguishable from native German speakers. This strongly suggests that the quality of the lexical representation is a prerequisite for the ability to use the information in an anticipatory fashion.

If lexical quality is indeed a determinant of predictive ability, one would expect that with increasing proficiency, L2 learners will have more high-quality lexical representations, which can be retrieved consistently and easily enough to be used predictively. Indeed, various studies have shown that predictive behavior increases with increasing proficiency (Chambers & Cooke, 2009; Dussias et al., 2013; Hopp, 2013) In addition, native speakers and L2 learners do not differ in predictive abilities when they have received the same amount exposure, e.g., when both are trained on novel words (Grüter et al., 2012). Future studies should more systematically probe the relation between the speed and consistency of lexical access, and the predictive use of that information in second language learners.

2.4 Task-induced processes and strategies

Predictive behavior is also affected by task-demands and the properties of other trials in the experiment. For instance, Farmer, Fine, and Jaeger (2011) report a reduction of the garden path effect for reduced relative clauses over the course of the experiment. They attribute this effect to the massive exposure to reduced relative clauses during the experiment, which are otherwise rather infrequent usages of the verbs used in the materials. Ferreira, Foucart, and Engelhardt (2013) show that predictive behavior in visual world eye-tracking paradigms is modulated by the number of elements in the display and the time the participant has to preview the display before hearing the instructions. Van Heugten, Dahan, Johnson, and Christophe (2012) demonstrated that the inclusion of gender mismatches between the determiner and the noun led to a reduced reliance on gender information to predict the upcoming noun. In addition, if the contexts are highly predictive, native speakers may not show extended effects of prediction and may stop processing before the highly expected word appears in the input (Besson, Faita, Czternasty, & Kutas, 1997).

Predictive behavior is therefore affected by task-induced strategies and by the modulation of attention. Therefore, apparent differences in predictive processing between native and non-native speakers could also be attributed in part to each group's differences in sensitivity to particular aspects of the task. For instance, due to their different exposure to the language, native and non-native speakers may be primed differently by preceding trials. There is some evidence that language users are more sensitive to structures that are less frequent (Jaeger & Snider, 2013). Since L2 learners may have different frequency biases than native speakers have (see Section 2.1), learners may be primed by the experimental materials in a different way than native speakers are. To give one hypothetical example, let's take the filled-gap paradigm used to investigate the processing of *wh*-constructions. In these paradigms, the *wh*-phrase or relativizer is the object of a stranded preposition downstream (e.g., *that is the magazine that the hairdresser read the article […] about…* Felser et al., 2012); the actual direct object position is occupied by an overt noun phrase (*the article*). Native speakers show processing difficulty at or after the overt direct object, whereas L2 speakers may show smaller or delayed effects (e.g., Dallas, 2008; Felser et al., 2012).[5] Let's assume that the L2 English speakers in the study have a native language that does not allow preposition stranding, or allows preposition stranding only in marked contexts. Let's further assume that these L2 speakers carry over these biases from their native language, and that they have a strong bias to interpret a *wh*-phrase as a direct object. These L2 participants may therefore be extremely sensitive to preposition stranding in the experimental materials, since the structure is rather frequent in the experiment, but infrequent in the language they have encountered thus far. The initial occurrences in the experiment may therefore lead to a strong violation of the L2 learners' predictions that the *wh*-phrase is a direct object. As a result, they may adjust their expectations quickly over the course of the experiment such that they come to expect the *wh*-phrase to be the object of a stranded preposition, or at least, not the direct object. Averaging over trials, filled-gap effects may be reduced and/or delayed in the L2 speakers compared with native speakers for whom preposition stranding is less noticeable because it frequently occurs in the language they have been exposed to all their lives. Crucially, one also predicts a larger difference in the size of the filled-gap effect between the beginning and the end of the experiment in L2 speakers compared with L1 speakers. Future experiments should include analyses of the change of processing patterns over the course of the experiment (Farmer et al., 2011). This makes it possible to test to what extent differences between native and L2 speakers can be due to different sensitivity to the experimental manipulation, and to differences in the way the participants adjust their processing over the course of the experiment.

2.5 Other factors

Predictive behavior can be affected by a variety of factors. I have discussed frequency, the amount of competing information, accuracy and consistency of the lexical information retrieved, and task-induced processes and strategies. This list is by no means exhaustive. For instance, motivation and emotional state can modulate the degree to which information is predicted, and the extent to which the prediction error is used to adjust expectations, and hence adapt to the context and learn (Clark, 2013). Predictive processing is also likely to be affected by the cognitive resources available and by cognitive control (Slevc & Novick, 2013, but see Otten & Van Berkum, 2009). For instance, the ability to inhibit irrelevant information may lead to a reduction of the number of alternatives and hence to stronger predictions concerning upcoming information. According to some studies, older adults have a reduced ability to inhibit information (Hasher, Lustig, & Zachs, 2008; Sommers & Danielson, 1999); this may account for the finding that elderly adults do not predict information to the same extent during sentence processing as young adults, especially when the context is weakly-constraining (DeLong, Groppe, Urbach, & Kutas, 2012; Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa & Kutas, 2002; Wlotko & Federmeier, 2012; Wlotko, Federmeier, & Kutas, 2012). Although in a speculative fashion, these factors can obviously extend to L2 processing. For instance, given the massive activation of competing information in L2 learners (see Section 2.2), one could predict that L2 learners would show more anticipatory processing if they are better at inhibiting irrelevant information.

2.6 Summary

Claiming that native speakers differ from L2 speakers in sentence processing in that native speakers actively predict and L2 learners do not runs counter to the available evidence, and begs the question of what determines when L2 speakers predict and when they do not. I have discussed several interdependent factors that may underlie differences in predictive processing and have shown how they could account for the differences in prediction processing that have been observed between native speakers and L2 learners, and among L2 users. There is therefore no need to stipulate a qualitative difference between native and non-native speakers; instead, the same mechanisms that affect prediction in native speakers can account for the apparent lack of prediction in some language learners.

3. Incorporating predictions in current proposals on L2 processing

The question of what underlies predictive behavior in L2 is not much different from the more commonly investigated question of what kind of information L2 speakers can use immediately during the on-line processing and integration of the current input. One slight difference is that models of predictive processing need to specify what information is available quickly and reliably enough to pre-activate what is yet to come in the input, and what exactly is pre-activated. Although most existing models of L2 processing do not explicitly include a predictive component, such a component can be easily incorporated.

3.1 Models of L2 processing

First, according to the Shallow Structure Hypothesis (SSH, Clahsen & Felser, 2006), non-native readers do not build detailed syntactic representations during processing. Instead, they employ heuristics based on the context and lexical information. Adjacent morphosyntactic relations (e.g., gender concord between nouns and determiners) are relatively easy for L2 learners to process; L2 speakers will have the most difficulty with syntactic movement and long-distance dependencies in which syntactic information is needed to construct the meaning of the sentence. The SSH can easily incorporate a predictive component: because L2 speakers rely on the context and on lexical information, and because adjacent dependencies are easy to process, this model can readily accommodate such findings as a) L2 speakers using semantic information to anticipate (Chambers & Cooke, 2009), and b) L2 speakers using gender information to predict upcoming nouns (Dussias et al., 2013; Hopp, 2013).

Exposure based models, such as constraint-based lexicalist models (MacDonald et al., 1994) and tuning models (Cuetos, Mitchell, & Corley, 1996) were primarily developed to account for native language processing, but have recently been applied to non-native processing (Dussias & Sagarra, 2007; Dussias & Cramer Scaltz, 2008; Frenck-Mestre & Pynte, 1997). These models assign an important role to frequency information: possible analyses for the input encountered are ranked on the basis of frequency information and can be seen as weighed predictions about what is to come next. Frequency information can be adjusted depending on errors and changes in experience. These models are therefore directly compatible with prediction-based accounts.

Another exposure-based model is the Competition Model (Bates & MacWhinney, 1981; MacWhinney & Bates, 1989). According to this model, language users employ certain cues to determine the meaning of the sentence (who did what to whom) such as animacy, word order, gender agreement, subject-verb

agreement, and case. The combination and relative weight of these cues differ per language. Language learners need to adjust the type and weight of the cues used in the second-language, and can experience interference from the settings in their native language. Although the Competition Model was mainly developed to account for differences in off-line processing preferences between languages and between native and non-native language users, the model can easily be applied to on-line anticipatory processing and account for the data discussed in this paper. If language users encounter a cue that is a reliable indicator of a certain syntactic role or function, they can use this to anticipate upcoming information. For instance, in German, an accusative noun phrase at the beginning of a clause is a good cue that the noun phrase is the direct object and that the next noun phrase is the subject. Indeed, native German speakers use such case information to anticipate the subject-agent of the clause (Kamide et al., 2003b). This model can therefore readily account for differences in the anticipatory use of cues between native and non-native speakers and for differences between non-native speakers of different language backgrounds (Dussias et al., 2013).

According to the Continuity Hypothesis and related processing deficit approaches advocated by, e.g. Hopp (Hopp, 2009, 2010) and McDonald (McDonald, 2006), differences between native and non-native speakers can be mainly attributed to resource deficits. Non-native speakers may be slower in accessing lexical items in the L2, or in applying parsing routines that are automatic in native speakers. When the sentence or task becomes more complex, non-native speakers may suffer from a shortage of resources and may show differences in processing with respect to native speakers. According to this account, non-native speakers can display the same anticipatory behavior as native speakers but may show differences if a) the task gets harder, and/or b) their access to lexical information is slower; and/or c) their parsing routines are less automatic. Supportive evidence comes from Hopp (2013), who found evidence that the anticipatory use of gender information is correlated with the speed of lexical access, in native as well as non-native speakers. In addition, native speakers under time pressure make errors similar to language-learners (Hopp, 2010; McDonald, 2006). A processing deficit account is also directly compatible with a prediction-based approach.

Current models of L2 processing can therefore easily be extended to include pre-activation of information before it appears in the input. Another essential component of a predictive processing approach is the use of the prediction errors (the difference between what is predicted and what is encountered in the input) to adjust future predictions. With the exception of some exposure-based models, this component is currently lacking from the major approaches.

3.2 What's new?

The idea of anticipatory processing is not new, and can easily be incorporated in existing models of native and non-native parsing. In fact, many recent presentations and publications propose to reinterpret existing data on e.g., garden pathing and semantic illusions in terms of predictive processing (e.g., Gibson, Steams, Bergen, Eddy, & Fedorenko, 2013; Kuperberg, 2013). So, what is the advantage of rethinking both native and non-native processing in terms of anticipatory processing? Re-interpreting processing in terms of anticipatory processing unifies the idea of parsing, learning, and exposure: during processing, predictions are made as to what is most likely to come in the input. These predictions are based on the frequency with which certain combinations of items have been encountered during previous exposure. If a prediction is not borne out, the error is used to adjust the frequency information stored about the words and structures used. In this approach, parsing, exposure and learning cannot be seen as isolated from the other, and should not be studied independently of each other.[6]

Empirically, what distinguishes anticipatory approaches from other accounts of native and non-native sentence processing, however, is that anticipatory approaches imply that the effects of an experimental manipulation can be observed *before* the critical word appears in the input, and that lexical quality, experimental context, and other factors discussed above will have an effect by this point. In addition, the nature of the prediction will have consequences for the recognition and integration of the actual critical word; hence, a relation is expected between processes observed before and after the presentation of the critical word (Molinaro et al., 2013). Furthermore, if predictions are adjusted based on the discrepancy between what is predicted and what is observed, the nature of the preceding trials in the experiment is expected to have a large effect on processing of a subsequent trial. This is especially the case when a strongly expected continuation is contradicted by the input: This contradiction may lead to a large adjustment of the information stored and a large modification of the processing of new trials relative to situations in which expectations are not violated in preceding trials (Jaeger & Snider, 2013). Again, native speakers and L2 learners are not expected to differ in these mechanisms, but rather in the nature and retrieval of the information used by these mechanisms.

4. Concluding remarks and future research

Rather than stipulating qualitative differences between L2 and native processing, research should aim to determine what drives the apparent differences between

language learners and native speakers. In the discussion above I have identified several factors that underlie apparent differences in predictive sentence processing between native and L2 speakers. Future research on anticipatory processing in L2 learners should take into consideration the nature of the L2 learner's input and exposure, combining corpus and on-line studies. In addition, non-circular tasks should be used to assess the accuracy and stability of the L2 speaker's lexical representations and how these relate to predictive behavior. Longitudinal studies, or at least multiple cross-sectional studies, need to be conducted to examine a) how lexical quality, frequency information stored, and the anticipatory use of information during on-line processing change as the L2 speaker becomes more proficient, b) how individuals differ, and c) what the predictors of efficient learning are. Moreover, task properties need to be taken into consideration, and participants need to be monitored as to their sensitivity to task-induced manipulations and changes in strategies over the course of the experiment. Finally, comparable studies need to be conducted on native speakers, in order to determine whether all differences between L2 and native speakers can be reduced to exposure and the other factors discussed above, or whether there is still something inherently different in how native and non-native speakers anticipate information during on-line sentence processing.

Notes

* The author was supported by a visiting scholars grant from the Netherlands Organization for Scientific Research (NWO 040.11.367). The author would like to thank Wind Cowles, Ana de Prada Perez, Jason Rothman, Steffi Wulff, Frank Wijnen, Alex Romero, George Collins, and the LAB reviewers for comments.

1. I will use the terms 'prediction', 'anticipation' and 'expectation' interchangeably. Others, e.g., Van Petten and Luka (2012), reserve the term 'prediction' for the expectation of a specific lexical item.

2. It is still controversial what exactly leads to processing difficulty (see Levy & Gibson, 2013). According to one view, the sentence will be harder to process when there are more unsatisfied predictions maintained over a longer distance (Gibson, 1998, 2000). According to another view, processing difficulty decreases when new input prunes the number of options (Levy, 2008). These two views make different predictions concerning the point at which difficulty occurs in sentence processing, and whether difficulty increases or decreases as the dependency gets longer. I will refrain from discussing this issue further here.

3. Another factor that may affect the strength of predictions is the type of information predicted. During sentence processing, semantic information may be more susceptible to decay, compared with syntactic category information, and hence will less likely be used predictively when the dependency gets longer (Wagers & Phillips, 2013). In addition, semantic information

may need some time to gain enough strength to steer predictions (Chow, Phillips, & Wang, 2012). Because only a few studies directly compare the predictive strength of various kinds of information, and, moreover, given that there are no studies using paradigms that probe processing before the predicted word or category is presented, I will not discuss this issue further.

4. The notion of "quality", i.e., stable and accurate information, can be expanded to include syntactic rules and information structure, which abstract away from specific lexical items.

5. I acknowledge that the hypothetical strategic differences sketched do not account for the entire data pattern observed in Felser et al. (2012). When the manipulation was semantic rather than syntactic, the L2 learners showed earlier implausibility effects than native speakers, suggesting that the L2 speakers did try to semantically (and systematically) combine the relative clause verb with the head of the relative clause (but see, Omaki & Schulz, 2011, for L2 speakers showing a slight delay in semantic effects relative to native speakers in a similar paradigm).

6. Predictive processing does not account for all processing phenomena. For instance, in case of backwards dependencies such as reflexives, the presence of a dependency can typically not be anticipated. Differences observed in the processing of such dependencies between native speakers and L2 learners (Felser & Cunnings, 2012) can therefore not be attributed to differences in predictive processing, although the same factors that underlie predictive processing (frequency, lexical quality, etc.) are likely to affect the processing of backward dependencies as well.

References

Abney, S. P., & Johnson, M. (1991). Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research*, *20*(3), 233–250. DOI: 10.1007/BF01067217

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

Altmann, G. T. M., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33* (4), 583–609. DOI: 10.1111/j.1551-6709.2009.01022.x

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, *364*(1521), 1235–1243. DOI: 10.1098/rstb.2008.0310

Bates, E., & MacWhinney, B. (1981). Second language acquisition from a functionalist perspective: Pragmatic, semantic and perceptual strategies. In H. Winitz (Ed.), *Annals of the New York Academy of Sciences conference on native and foreign language acquisition* (pp. 190–214). New York: New York Academy of Sciences.

Besson, M., Faita, F., Czternasty, C., & Kutas, M. (1997). What's in a pause: event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology*, *46*(1), 3–23. DOI: 10.1016/S0301-0511(96)05215-5

Boland, J. E. (2005). Visual arguments. *Cognition*, *95*(3), 237–274. DOI: 10.1016/j.cognition.2004.01.008

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, *34*, 774–806. DOI: 10.1006/jmla.1995.1034

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

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

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. DOI: 10.1037/0033-295X.113.2.234

Chow, W.-Y., Phillips, C., & Wang, S. (2012). Turning the 'Dumb N400' into the 'Smart N400': What role-reversed sentences tell us about the time course of predictions. Talk presented at *CUNY Conference on Human Sentence Processing*, New York, NY.

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

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. DOI: 10.1017/S0140525X12000477

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 94–128). Cambridge: Cambridge University Press.

Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. E. García-Albea, & N. Sebastián-Gallés (Eds.), *Language processing in Spanish* (pp. 145–187). Mahwah, NJ: Lawrence Erlbaum Associates.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language*, *42*(4), 465–480. DOI: 10.1006/jmla.1999.2688

Dallas, A. C. (2008). *Influences of verbal properties on second-language filler-gap resolution: A cross-methodological study*. Unpublished doctoral dissertation, University of Florida.

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 33–40). Somerville, MA: Cascadilla Proceedings Project.

DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language*, *121*(3), 226–239. DOI: 10.1016/j.bandl.2012.02.006

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

Dickey, M. W., Choy, J. J., & Thompson, C. K. (2007). Real-time comprehension of wh-movement in aphasia: Evidence from eyetracking while listening. *Brain and Language*, *100*(1), 1–22. DOI: 10.1016/j.bandl.2006.06.004

Dickey, M. W., & Thompson, C. K. (2009). Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking. *Journal of Neurolinguistics*, *22*(6), 563–583. DOI: 10.1016/j.jneuroling.2009.06.004

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496–518. DOI: 10.1006/jmla.1999.2654

Dikker, S., & Pylkkänen, L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language*, *127*, 55–64. DOI: 10.1016/j.bandl.2012.08.004

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, *21*(5), 629–634. DOI: 10.1177/0956797610367751

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*(3), 293–321. DOI: 10.1016/j.cognition.2008.09.008

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish-English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologia*, *128*, 501–513. DOI: 10.1016/j.actpsy.2007.09.004

Dussias, P. E., Marful, A., Gerfen, C., & Bajo, M. T. (2010). Usage frequencies of complement-taking verbs in Spanish and English: Data from Spanish monolinguals and Spanish—English bilinguals. *Behavior Research Methods*, *42*(4), 1004–1011. DOI: 10.3758/BRM.42.4.1004

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. Bilingualism: *Language and Cognition*, *10*(1), 101–116. DOI: 10.1017/S1366728906002847

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, *35*, 353–387. DOI: 10.1017/S0272263112000915

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, *7*(2–3), 195–225.

Farmer, T. A., Fine, A. B., & Jaeger, T. F. (2011). Implicit context-specific learning leads to rapid shifts in syntactic expectations. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual meeting of the Cognitive Science Society* (pp. 2055–2060). Austin, TX: Cognitive Science Society

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, *115*(3), 149–161. DOI: 10.1016/j.bandl.2010.07.006

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(02), 133–146. DOI: 10.1111/1469-8986.3920133

Felser, C., & Cunnings, I. (2012). Processing reflexives in a second language: The timing of structural and discourse-level constraints. *Applied Psycholinguistics*, *33*(03), 571–603. DOI: 10.1017/S0142716411000488

Felser, C., Cunnings, I., Batterham, C., & Clahsen, H. (2012). The timing of island effects in nonnative sentence processing. *Studies in Second Language Acquisition*, *34*, 67–98. DOI: 10.1017/S0272263111000507

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, *69*(3), 165–182. DOI: 10.1016/j.jml.2013.06.001

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior*, *18*(1), 1–20. DOI: 10.1016/S0022-5371(79)90534-6

Flett, S., Branigan, H. P., & Pickering, M. J. (2013). Are non-native structural preferences affected by native language preferences? *Bilingualism: Language and Cognition*, *First View*, 1–10. DOI: 10.1017/S1366728912000594

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading second and native languages. *The Quarterly Journal of Experimental Psychology A*, *50*(1), 119–148. DOI: 10.1080/027249897392251

Garnsey, S. M., Pearlmutter, N. J., Myers, E. M., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93. DOI: 10.1006/jmla.1997.2512

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, *68*(1), 1–76. DOI: 10.1016/S0010-0277(98)00034-1

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, Language, Brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., Steams, L., Bergen, L., Eddy, M., & Fedorenko, E. (2013). The P600 indexes rational error correction within a noisy-channel model of human communication. Talk presented at The 26th Annual CUNY Conference on Human Sentence Processing, Columbia, SC.

Gollan, T. H., & Acenas, L.-A. R. (2004). What Is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246–269. DOI: 10.1037/0278-7393.30.1.246

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(7), 87–814. DOI: 10.1016/j.jml.2007.07.001

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*, 1220–1234. DOI: 10.3758/BF03193224

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The Frequency-Lag Hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. DOI: 10.1037/a0022256

Gries, S. T., & Wulff, S. (2005). Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics*, *3*, 182–200. DOI: 10.1075/arcl.3.10gri

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, *28*(2), 191–215. DOI: 10.1177/0267658312437990

Grüter, T., & Rohde, H. (2013). *L2 processing is affected by RAGE: Evidence from reference resolution*. Talk presented at the 12th conference on Generative Approaches to Second Language Acquisition (GASLA), University of Florida, FL.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistics Research*, *30*, 251–266. DOI: 10.1023/A:1010490917575

Hahne, A., & Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials . *Bilingualism: Language and Cognition*, *4*(2), 123–141 DOI: 10.1017/S1366728901000232.

Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, *30*(4), 609–642. DOI: 10.1207/s15516709cog0000_64

Hasher, L., Lustig, C., & Zachs, R. (2008). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in*

working memory (pp. 227–249). New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780195168648.003.0009

Hopp, H. (2009). The syntax-discourse interface in near-native L2 acquisition: Off-line and on-line performance. Bilingualism: *Language and Cognition*, *12*(4), 463–483. DOI: 10.1017/S1366728909990253

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

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

Jacob, G. (2009). *The role of native language in second-language syntactic processing*. Unpublished doctoral dissertation, University of Dundee.

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83. DOI: 10.1016/j.cognition.2012.10.013

Kaan, E., Dallas, A. C., & Wijnen, F. (2010). Syntactic predictions in second-language sentence processing. In J.-W. Zwart, & M. de Vries (Eds.), *Structure preserved. Festschrift in the honor of Jan Koster* (pp. 207–213). Amsterdam: John Benjamins.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003a). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language*, *49*, 133–159. DOI: 10.1016/S0749-596X(03)00023-8

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003b). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*(1), 37–55. DOI: 10.1023/A:1021933015362

Kim, A. E., & Gilley, P. M. (2013). Neural mechanisms of rapid sensitivity to syntactic anomaly. *Frontiers in Psychology*, *4*, 45. DOI: 10.3389/fpsyg.2013.00045.

Kuperberg, G. R. (2013). Predicting meaning: What the brain tells us about the architecture of language comprehension. Talk presented at the *26th Annual CUNY Conference on Human Sentence Processing*, Columbia, SC.

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207). New York, NY: Oxford University Press.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, *98*(1), 74–88. DOI: 10.1016/j.bandl.2006.02.003

Lee, E.-K., Lu, D. H.-Y., & Garnsey, S. M. (2013). L1 word order and sensitivity to verb bias in L2 processing. *Bilingualism: Language and Cognition*, *16*(4), 761–775. DOI: 10.1017/S1366728912000776

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177. DOI: 10.1016/j.cognition.2007.05.006

Levy, R., & Gibson, E. (2013). Surprisal, the PDC, and the primary locus of processing difficulty in relative clauses. *Frontiers in Psychology*, *4*, 229. DOI: 10.3389/fpsyg.2013.00229

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, *63*(4), 447–464. DOI: 10.1016/j.jml.2010.07.003

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226. DOI: 10.3389/fpsyg.2013.00226..

MacDonald, M. C., Pearlmuttter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703. DOI: 10.1037/0033-295X.101.4.676

MacWhinney, B., & Bates, E. (1989). *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*, 53–78. DOI: 10.1017/S0272263105050035

Martin, C., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588. DOI: 10.1016/j.jml.2013.08.001

Martinez-Garcia, M. T., & Wulff, S. (2012). Not wrong, yet not quite right: Spanish ESL students' use of gerundial and infinitival complementation. *International Journal of Applied Linguistics*, *22* (2), 225–244. DOI: 10.1111/j.1473-4192.2012.00310.x

McDonald, J. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*, 381–401. DOI: 10.1016/j.jml.2006.06.006

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eyemovements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5* (1), 1–10.

Molinaro, N., Barraza, P., & Carreiras, M. (2013). Long-range neural synchronization supports fast and efficient reading: EEG correlates of processing expected words in sentences. *NeuroImage*, *72*(0), 120–132. DOI: 10.1016/j.neuroimage.2013.01.031

Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition*, *125* (2), 317–323. DOI: 10.1016/j.cognition.2012.07.016

Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constaints in second-language sentence processing. *Studies in Second Language Acquisition*, *33*, 563–588. DOI: 10.1017/S0272263111000313

Otten, M., & Van Berkum, J. J. A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, *1291*, 92–101. DOI: 10.1016/j.brainres.2009.07.042

Perfetti, C. (2007). Reading ablility: lexical quality to comprehension. *Scientific studies of reading*, *11*(4), 357–383. DOI: 10.1080/10888430701530730

Perfetti, C., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection. Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, D.C.: American Psychological Association. DOI: 10.1037/10459-004

Pickering, M. J., & Garrod, S. (2011). The use of prediction to drive alignment in dialogue. In G. R. Semin, & G. Echterhoff (Eds.), *Grounding sociality. Neurons, mind and culture* (pp. 175–191). New York: Taylor & Francis.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. DOI: 10.1017/S0140525X12001495

Schilk, M. (2011). *Structural nativization in Indian English lexicogrammar*. Amsterdam: John Benjamins. DOI: 10.1075/scl.46

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10*, 209–241. DOI: 10.1515/iral.1972.10.1-4.209

Siyanova, A., & Conklin, K. (2008). Processing of binomials in sentence contexts by monolinguals and proficient bilinguals: Evidence from response times and eye-tracking. Talk presented at *21st Annual CUNY conference on Human Sentence Processing*, Chapel Hill, NC.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, *64*(429–458). DOI: 10.3138/cmlr.64.3.429

Slevc, L. R., & Novick, J. M. (2013). Memory and cognitive control in an integrated theory of language processing. *Behavioral and Brain Sciences*, *36*(4), 373–374. DOI: 10.1017/S0140525X12002683

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults. *Psychology and Aging*, *14*, 458–472. DOI: 10.1037/0882-7974.14.3.458

Spivey-Knowlton, M. J., Tanenhaus, M. K., Eberhard, K., & Sedivy, J. (1998). Integration of visuospatial and linguistic information: Language comprehension in real-time and real-space. In P. Olivier, & K. Gapp (Eds.), *Representation and processing of spatial expressions* (pp. 201–214). Hillsdale, NJ: Erlbaum.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10* (3), 281–284. DOI: 10.1111/1467-9280.00151

Stowe, L. A., Tanenhaus, M. K., & Carlson, G. N. (1991). Filling gaps on-line: use of lexical and semantic information in sentence processing. *Language and Speech*, *34*, 319–340.

Sussman, R. S., & Sedivy, J. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes*, *18*(2), 143–163. DOI: 10.1080/01690960143000498

Tabor, W., & Tanenhaus, M. K. (1999). Dynamical models of sentence processing. *Cognitive Science*, *23*(4), 491–515. DOI: 10.1207/s15516709cog2304_5

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 443–467. DOI: 10.1037/0278-7393.31.3.443

Van Hell, J., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, *9*(4), 780–789. DOI: 10.3758/BF03196335

van Heugten, M., Dahan, D., Johnson, E. K., & Christophe, A. (2012). Accommodating syntactic violations during online speech perception. Poster presented at *the 25th Annual CUNY Conference on Human Sentence Processing*, New York, NY.

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

Wagers, M. W., & Phillips, C. (2013). Going the distance: Memory and decision-making in active dependency construction. *Quarterly Journal of Experimental Psychology*. DOI: 10.1080/17470218.2013.858363.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory & Language*, *50*, 1–25. DOI: 10.1016/S0749-596X(03)00105-0

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, *99*(2), B63-B72. DOI: 10.1016/j.cognition.2005.07.001

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, *39*(3), 483–508. DOI: 10.1016/S0010-9452(08)70260-0

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

Wilson, M. P., & Garnsey, S. M. (2009). Making simple sentences hard: Verb bias effects in simple direct object sentences. *Journal of Memory and Language*, *60*(3), 368–392. DOI: 10.1016/j.jml.2008.09.005

Wlotko, E. W., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology*, *49*, 770–785. DOI: 10.1111/j.1469-8986.2012.01366.x

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*, *27*(4), 975–988. DOI: 10.1037/a0029206

Wulff, S., Lester, N., & Martinez-Garcia, M. T. (in press). That-variation in German and Spanish L2 English. *Language and Cognition*.

Author's address

Edith Kaan, Ph.D.
Department of Linguistics
University of Florida
Office: Turlington Hall 4127
Mailing address: Box 115454
Gainesville, FL 32611
USA

kaan@ufl.edu