# Predicting upcoming information in native-language and non-native-language auditory word recognition*

ASTER DIJKGRAAF
*Department of Experimental Psychology, Ghent University*
ROBERT J. HARTSUIKER
*Department of Experimental Psychology, Ghent University*
WOUTER DUYCK
*Department of Experimental Psychology, Ghent University*

*Monolingual listeners continuously predict upcoming information. Here, we tested whether predictive language processing occurs to the same extent when bilinguals listen to their native language vs. a non-native language. Additionally, we tested whether bilinguals use prediction to the same extent as monolinguals. Dutch–English bilinguals and English monolinguals listened to constraining and neutral sentences in Dutch (bilinguals only) and in English, and viewed target and distractor pictures on a display while their eye movements were measured. There was a bias of fixations towards the target object in the constraining condition, relative to the neutral condition, before information from the target word could affect fixations. This prediction effect occurred to the same extent in native processing by bilinguals and monolinguals, but also in non-native processing. This indicates that unbalanced, proficient bilinguals can quickly use semantic information during listening to predict upcoming referents to the same extent in both of their languages.*

In monolingual (native) language comprehension, people continuously generate predictions about upcoming input (e.g., Altmann & Kamide, 1999; Boland, 2005; DeLong, Urbach & Kutas, 2005). In a seminal paper, Altmann and Kamide (1999) studied prediction in auditory language comprehension using a visual world paradigm. Participants listened to sentences such as *The boy will eat the cake* or *The boy will move the cake*. Eye movements were recorded while participants viewed a visual scene with four objects that could all be moved, but in which only one object (the cake) was edible. When participants heard the verb *eat*, participants initiated fixations to the picture of the cake more often before the onset of the word *cake* than after hearing the verb *move*. Altmann and Kamide concluded that the sentence context pre-activated the representation of the target word. Various recent models of monolingual sentence comprehension have now incorporated predictive processing (e.g., Levy, 2008; MacDonald, 2013; Pickering & Garrod, 2013).

Using context information to generate predictions is fundamental in efficient language processing: it can speed up processing, solve ambiguities, and help the listener determine when to start an overt response in a dialogue (Kutas, DeLong & Smith, 2011; Van Berkum, 2010). These facilitatory functions could be particularly relevant in L2 comprehension, which is often considered to be slower, less accurate, and more resource-consuming than L1 processing (Cook, 1997; Hahne, 2001; Weber & Broersma, 2012). On the other hand, L2 processing difficulty may also impede efficient prediction during language comprehension. However, in spite of its possible increased importance, there is very little research about whether bilinguals predict input in their L2 as native speakers do in L1 or whether L2 words and their features are just integrated incrementally when they are encountered in the input rather than before.

In a recent review, Kaan (2014) suggested that predictive processing in L2 is not inherently different from predictive processing in L1, but that it may be modulated by factors associated with non-native comprehension. For example, it is often assumed that predictions are based on statistical regularities extracted from the input throughout a person's lifetime (e.g., Bar, 2007; MacDonald, 2013). However, information stored in memory about how often a word tends to occur in a certain context (e.g., an edible object following the verb *eat*) may be different in L2 and L1 because the L2 has usually been practiced less (Gollan, Montoya, Cera & Sandoval, 2008) and in different settings (e.g., native learning versus classroom learning). Less or different input in L2 may affect the content and strength of predictions. Importantly, if L2 is practiced less than L1, representations of lexical form,

Address for correspondence:
A. Dijkgraaf, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium
aster.dijkgraaf@ugent.be

meaning and use as well as the links between them may be less consistent and less accurate in L2 (Gollan et al., 2008). Weaker representations may lead to less efficient retrieval. And less efficient retrieval of lexical form or semantic associations may in turn lead to slower, less accurate or weaker predictions. Likewise, because bilinguals divide language use between L1 and L2, and therefore also have less L1 practice, L1 processing too may be different for monolinguals and bilinguals. If inconsistency of lexical representations indeed affects prediction skill during comprehension, then prediction skill is expected to increase with increased consistency of representations. This implies that predictive processing in L2 should become more native-like as L2 proficiency increases.

Furthermore, lexical competition is increased in L2 processing because of simultaneous activation of L1 words and because L2 speakers often misperceive phonemes, thereby increasing the number of words perceived as similar (Lagrou, Hartsuiker & Duyck, 2013a; Weber & Cutler, 2004). Increased competition can cause a delay in the selection of a predicted word, as well as in processing the context information used to generate a prediction. Finally, a number of other factors are thought to modulate prediction in monolingual language processing, such as resource limitations, emotional state and cognitive control. Kaan (2014) suggests that the effect of each of these factors may in turn interact with processing language (native or non-native), so that L2 data is required to evaluate the generalizability of each demonstration of prediction in monolingual language processing.

Some studies reveal effects of semantic context on target word recognition (Chambers & Cooke, 2009; FitzPatrick & Indefrey, 2007; Lagrou, Hartsuiker & Duyck, 2013b) in L2 processing. However, effects found at presentation of the target word do not allow us to distinguish facilitation of semantic integration from semantic prediction. A constraining sentence context may facilitate word integration upon presentation of the word in L2 processing, but whether or not bilinguals actively predict information, online and during sentence processing, to the same extent in L1 and L2, remains unclear.

### Prediction in L2 Reading

In a study in the visual domain by Martin, Thierry, Kuipers, Boutonnet, Foucart and Costa (2013), native speakers of English and late Spanish–English bilinguals read sentences in English with predictable or less predictable sentence-final nouns. Event-related potentials were measured at the article preceding the sentence-final noun. The article was always congruent with the final noun, but not with the expected noun (e.g., *Since it is raining, it is better to go out with an*

*umbrella [EXPECTED]/ a raincoat [UNEXPECTED]*). If participants indeed predicted *umbrella*, a semantic anomaly effect should be elicited by the article *a* relative to *an*, because *a* is incongruent with *umbrella*. Thus, the target for prediction is the lexical form and the congruent article. The target is predicted based on semantic information from the sentence context. Martin et al. indeed found an N400-effect for the incongruent article for L1 readers, but not for L2 readers. The lack of an effect on the article was taken to indicate that L2 readers did not predict the target word (at least not as efficiently as L1 readers). For the target noun, the authors did find a significant N400-effect in central and parietal regions in both L1 and L2 readers, but the effect was significantly larger in L1 than in L2 readers. The N400-effect on the noun showed that even though the participants reading in L2 did not predict upcoming input, integration of a target word in the sentence was still easier if the sentence was constraining.

The lack of a prediction effect on the article in L2 comprehension in Martin et al.'s study (2013) may have resulted from the particular manipulation used. In particular, the lexical prediction effect was measured on the basis of the congruency of an article (*a/an*) with the predicted word. The particular phonological agreement rule manipulated does not exist in the bilingual participants' L1. Martin et al. (2013) tested whether a group of intermediate L2 proficient participants, not participating in their experiment, knew the phonological article-noun agreement rule. Both an online and an offline test showed that intermediate L2 proficient participants were sensitive to the agreement rule. However, the intermediate L2 proficient group actually participating in the experiment may not have been able to apply the rule quickly enough for a prediction-incongruent determiner to modulate the N400 effect.

Therefore, in a second study in the visual domain, Foucart, Martin, Moreno and Costa (2014), used a similar sentence reading paradigm but measured the prediction effect by manipulating prediction congruency of the determiners' gender in Spanish sentences (e.g*., The pirate had the secret map, but he never found the [masc] treasure [EXPECTED]/ the [fem] cave [UNEXPECTED] he was looking for.*). As in Martin et al., the target for prediction was the lexical form and the congruent article. The target is predicted based on semantic information from the sentence context. However, in this study the gender agreement rule between the target article and noun existed both in the late bilingual participants' L1 (French) and L2 (Spanish). Here, the authors found an effect of congruency of the article and the predicted noun on the N400 elicited by the article both in L1 reading (by Spanish monolinguals and early Spanish–Catalan bilinguals) and in L2 reading (by late French–Spanish bilinguals), although the effect lasted for a shorter time in the late bilingual group. The

results demonstrate that bilinguals reading in L2 can use semantic information from the sentence context to predict upcoming words and their gender. Foucart et al. suggested that the similarity between the article-noun agreement rule in late bilingual participants' L1 and L2 may have made it easier for the participants to generate a prediction in time. In addition, half of the expected nouns included in the experiment were cognates, possibly adding to the facilitatory effect. The two studies described above show that bilinguals can predict lexical information in sentence reading, but that whether or not prediction occurs may depend on L1 and L2 language similarity.

### Prediction in L2 listening

Both studies described above were conducted in the visual domain, but predictive language processing may well be more challenging in the auditory modality. For instance, the fact that auditory input unfolds over time, unlike written input, may make prediction more relevant because the listener cannot return to prior input or influence input rate, unlike in reading[1]. Predictive processing may also be more difficult in the auditory modality than in the visual modality for bilinguals because of increased cross-language co-activation due to misperceptions and misrepresentation in listening (Weber & Cutler, 2004).

Foucart, Ruiz-Tada and Costa (2015) tested prediction in the auditory modality using an EEG paradigm similar to Foucart et al. (2014). Again, the target for prediction was the lexical form with the congruent article, and predictions were based on semantic information from the sentence context. The authors found that bilinguals listening in L2 are able to predict upcoming words based on sentence context. The participants in this study were all bilingual and they were only tested in their L2. Therefore, no direct comparison could be made between the size of the effect in L1 and L2 in bilinguals, or between the size of the effect in monolinguals (L1) and bilinguals (L1 or L2).

Visual world paradigm studies on prediction in L2 auditory processing have mainly focused on prediction based on morphosyntactic information. In a visual world experiment, Hopp (2013) investigated whether German native and English–German bilingual listeners would show predictive looks to target objects whose gender agreed with an article in the auditory signal. Like native listeners, English–German bilinguals listening in L2 were more likely to look at the target objects whose gender agreed with an afore-mentioned article before the onset of the target object in the auditory signal, but only in the bilinguals who had native-like mastery of gender assignment.

Hopp (2015) used a visual world paradigm to investigate whether English–German bilinguals integrate morphosyntactic information and verb semantics to generate predictions about upcoming semantic input during L2 auditory comprehension. In this experiment, picture displays including three possible actors and a control object were paired with an SVO (e.g., *The$_{NOM}$ wolf kills soon the$_{ACC}$ deer*) or an OVS (e.g., *The$_{ACC}$ wolf kills soon the$_{NOM}$ hunter*) sentence in German. Native listeners were more likely to look at expected patients (e.g., *the deer)* before the onset of the second NP in SVO sentences and at expected agents (e.g., *the hunter*) in OVS sentences. English–German bilinguals, on the other hand, were more likely to fixate patients before the onset of the second NP, independently of the case marking (nominative or accusative) of the first NP. Hopp concluded that there was an effect of semantic prediction in L2 based on information extracted at the verb, but that case information did not modulate predictions like in L1 listeners. Bilingual participants seemed unable to apply an L2 agreement rule not present in their L1 on the fly, or at least not quickly enough to support prediction. Hopp's findings are in line with recent findings of Mitsugi and Macwhinney (2016), who demonstrated that L1 English learners of Japanese with good offline knowledge of the Japanese case-marking system were unable to employ this knowledge online in order to generate predictions in a visual world eye-tracking experiment.

Dussias, Valdés Kroff, Guzzardo Tamargo and Gerfen (2013) also focused on prediction based on morphosyntactic information, specifically, prediction based on article-noun gender agreement. A group of English–Spanish bilinguals (high and low proficiency), Italian–Spanish bilinguals and Spanish monolinguals saw a display with two pictures of items with the same or different grammatical gender. While looking at the display, they heard a sentence with an article that either agreed with the gender of one of the two items in the display, or with both. Spanish monolinguals looked at the target picture sooner in the different gender condition (when the article was a cue) than in the same gender condition. Highly proficient English–Spanish bilinguals, but not low proficient English–Spanish bilinguals, also looked at the target picture earlier in the different gender condition. Unlike the low proficient English–Spanish bilinguals, low proficient Italian–Spanish bilinguals looked at the target picture significantly earlier in the different gender condition, but only when the target item was feminine. Dussias et al.'s results suggest that highly proficient bilinguals use gender cues to anticipate information as monolinguals do, whereas low proficient bilinguals do not, unless their native language has a similar article-noun gender agreement system. Even though the effects Dussias et al. found for monolinguals and highly proficient bilinguals

---

[1] In Martin et al. (2013), the first half of each stimulus sentence was presented on the screen as a whole. After pressing spacebar, one word was presented every 700 ms.

are likely to be anticipatory in nature, given their time course, the authors do not explicitly distinguish between effects anticipation and facilitation of integration.

These recent visual world studies on prediction in L2 listening reveal that it is especially difficult for bilinguals to process morphosyntactic features quickly enough to use them as a cue to generate predictions in L2. However, it remained unclear whether bilinguals also have difficulty anticipating semantic information in L2 processing, which would always lead to weaker L2 prediction effects, or, whether they selectively have difficulty applying language-specific and, difficult, grammatical rules quickly enough during predictive processing. Hopp (2015) explicitly distinguishes prediction based on verb semantics and prediction based on case-marking. However, as Hopp proposes, the significant effect of prediction based on verb semantics (predictive looks to the patient object in both SVO and OVS sentences) in L2 listening can be interpreted in two ways: either the L2 listeners used semantic information extracted at the verb to guide predictive looks towards the most plausible sentence object in the picture display (the patient), or, on the basis of the first NP, fixations were directed to a plausible patient object, regardless of verb semantics. Therefore, it remains unclear whether bilinguals are able to use verb semantics to guide their predictions during non-native sentence comprehension as they do in L1.

Koehne and Crocker (2015) provided evidence that language learners are able to use semantic restrictions at the verb to predict upcoming referents. Participants learned novel, artificial verb, subject (*man* and *woman*) and object names by exposure to verbs with visual context, followed by exposure to nouns in SVO sentence context, in a visual world paradigm. Anticipatory eye-movements to the sentence target objects were found during presentation of the constraining verb. As each verb type was combined with each subject type, the anticipatory eye-movements to the target object could not have been based on information extracted at the sentence subject alone. Koehne and Crocker show that people can use verb semantics to predict upcoming information in early language learning. However, instruction specifically stressed semantic processing of the sentences. Also, a limited number of artificial verbs (six at most) and objects (18 at most) were used in the study. These two factors may have greatly inflated predictive processing when compared to natural L2 language processing.

### Present Study

All previous studies on anticipating information in L2 listening have either focused on L2 listening alone, or they have compared a group of L2 listeners to a group of L1 listeners in a between-participants design. In the present experiment, Dutch–English bilinguals were tested in the native and non-native language. In addition, an English monolingual control group was tested in order to compare L1 with L2 listening in the same language (English) and L1 (English) listening by monolinguals with L1 (Dutch) listening by bilinguals. Comparing predictive processing within participants is important, as recent studies have shown effects of cognitive factors such as verbal fluency, vocabulary size (Rommers, Meyer & Huettig, 2015), working memory and processing speed (Huettig & Janse, 2016) on predictive language processing. There may also be factors inherent to bilingualism (and not L2 processing) that affect predictive processing. For example, bilinguals activate lexical information in both languages during L1 and L2 processing (e.g., Lagrou et al., 2013a). Bilinguals may therefore activate more information during language processing which in turn may slow down the prediction process. In addition, some authors suggest that bilinguals have increased cognitive control abilities compared to monolinguals (Woumans, Ceuleers, Van der Linden, Szmalec & Duyck, 2015). Increased cognitive control may help suppress irrelevant information during predictive processing. For example, Zirnstein, Hell and Kroll (2015) recently found that that processing costs for unverified predictions were larger in low-control than in high-control bilingual participants. In this experiment, we will compare bilinguals listening to speech in L1 and L2 to eliminate effects of individual differences. As a control experiment, we will also compare bilinguals (L2) to monolinguals (L1) listening to the same language (English). Finally, to test whether there are any effects of speaker bi- or monolingualism on predictive language processing we will compare prediction effects in L1 processing in bilinguals (Dutch) to L1 processing in monolinguals (English).

Here, a visual world paradigm based on Altmann and Kamide's (1999) task was used. Participants listened to sentences such as *Mary knits a scarf* or *Mary loses a scarf*. Eye-movements were recorded while participants viewed a visual scene with four objects that could all be *lost* (neutral condition), but in which only one object (the scarf) was *knittable* (constraining condition). If participants predicted the target object in the constraining condition, this would result in a higher proportion of looks to the target object in the constraining condition than in the neutral condition before the onset of the target in the auditory stimulus. Based on Kaan (2014) we expected that bilinguals listening in L2 would not predict semantic properties of upcoming referents as fast and to the same extent as when listening in L1 because of modulating factors associated with L2 language processing, such as differences in stored statistical regularities and weaker, less accurate lexical representations. Further, we expected that bilingual participants listening in L1 would not predict semantic input to the same extent as monolinguals do in L1. This would be in line with the weaker links

Table 1. *Participants' mean scores on proficiency tests and mean ratings*

|  | Bilinguals Dutch (SD) | Bilinguals English) (SD) | Monolinguals English (SD) | p-value[c] bilinguals L1 vs. L2 | p-value[d] monolinguals L1 vs. bilinguals L1 | p-value[e] Monolinguals L1 vs. bilinguals L2 |
|---|---|---|---|---|---|---|
| Lextale[a] | 86.13 (5.54) | 78.50 (10.49) | 87.83 (7.97) | <0.001 | .34 | <0.001 |
| Rating speaking | 9.2 (0.75) | 7.3 (1.34) | 9.6 (0.72) | <0.001 | .03 | <0.001 |
| Rating listening | 9.3 (0.79) | 8.1 (0.73) | 9.5 (0.78) | <0.001 | .07 | <0.001 |
| Rating reading | 9.3 (0.66) | 8.0 (1.22) | 9.3 (0.92) | <0.001 | .44 | <0.001 |
| Mean ratings[b] | 9.3 (0.7) | 7.8 (0.9) | 9.5 (0.71) | <0.001 | .09 | <0.001 |

[a]Scores consist of percentage correct, corrected for unequal proportion of words and nonwords (Lemhöfer & Broersma, 2012).
[b]Score based on means of self-assessed ratings on a scale of 1 to 10 (1=very low, 10=perfect) of speaking, listening and reading.
[c]Reported p-values indicate significance levels of dependent samples t-tests between scores for Dutch and English in bilinguals. Df of all t-tests = 29.
[d]Reported p-values indicate significance levels of independent samples t-tests between scores for bilinguals in Dutch and monolinguals in English. Df of all t-tests = 29.
[e]Reported p-values indicate significance levels of independent samples t-tests between scores for bilinguals and monolinguals in English. Df of all t-tests = 29.

hypothesis (Gollan et al., 2008) of bilingual language processing. This hypothesis states that bilinguals divide language use between L1 and L2, and therefore have less practice in each of their languages. Less practice in each language should lead to weaker links between semantics and phonology in bilinguals than in monolinguals and thereby to slower lexical access. In turn, these weaker links may result in slower or weaker predictive processing.

As opposed to previous studies on predictive processing in the non-native language, we opted for a design in which no language-specific agreement rule needed to be applied by the participants on the fly in order to measure the prediction effect or in order for the participant to make a prediction. This way, if we find an attenuation of the prediction effect in non-native listening, it cannot be attributed to difficulty applying a non-native agreement rule on the fly.

Finally, previous studies have suggested that predicting upcoming information during language processing serves as a learning mechanism (Dell & Chang, 2013; Koehne & Crocker, 2015; Mani & Huettig, 2012). For example, Mani and Huettig (2012) found a significant positive correlation between prediction skill and expressive vocabulary in children. We therefore expect that prediction effects are modulated by language proficiency, so that bilinguals with a higher proficiency score show a stronger prediction effect than bilinguals with a lower proficiency score.

dominant and most proficient language in the LEAP-Q questionnaire (Marian, Blumenfeld & Kaushanskaya, 2007), and English as their second (25 participants) or third (5 participants) language. Belgian and Dutch students typically start to learn English at age ten or eleven in school, and their English proficiency is relatively high because of regular input from popular media and study books. None of the participants had immersion experience in an English-dominant environment. On average the participants reported that they were exposed to English 17% of the time, versus 73% to Dutch. Besides knowledge of English and Dutch, twenty-eight participants had knowledge of French, and nineteen participants had knowledge of German. Fewer than six participants had knowledge of other languages such as Spanish, Turkish, Portuguese, Polish or Italian. To assess language proficiency in both languages, participants carried out the LexTALE vocabulary knowledge test (Lemhöfer & Broersma, 2012) and provided self-ratings. The LexTALE is an unspeeded 60-item lexical decision task. It is an indicator of word knowledge and general language proficiency (Lemhöfer & Broersma, 2012). The bilinguals' mean LexTALE scores and self-ratings are reported in Table 1. The LexTALE score and self-ratings show that the bilingual participants were more proficient in their native (Dutch) than in their non-native language (English).

## Methods

### *Participants*

### *Bilinguals*
Thirty native speakers of (Belgian or Netherlands) Dutch took part in the experiment (5 men and 25 women, mean age 24 years, range 20–41). They were recruited from the Ghent University participant database. All signed informed consent. All participants reported Dutch as their

### *Monolinguals*
Thirty monolingual native speakers of English participated in the experiment (4 men and 26 women, mean age 20 years, range 18–28). They were recruited from the University of Southampton participant database. All signed informed consent. The monolinguals' mean LexTALE scores and self-ratings are reported in Table 1. The LexTALE score shows that the bilingual and monolingual participants were matched on L1 proficiency.

Figure 1 Example Picture Display. The sentences belonging to this display were: Mary reads a letter and Mary steals a letter.

### Materials and design

Eighteen stimulus sets were created. Each set consisted of a four-picture display, two sentences in Dutch and their translation equivalents in English. One of the two sentences was constraining; the other sentence was neutral. In the constraining condition, only one of the objects in the display was appropriate after the verb, whereas all objects in the display were appropriate after the verb in the neutral condition (see Figure 1). Appendix A contains the constraining and neutral verbs as well as the objects in the display for each stimulus set[2].

Likewise, eighteen filler sets were created. Each set again consisted of a display with four pictures, two sentences in Dutch and their English translation equivalents. In the filler sets, sentences could apply to either no, or two or three objects in the display. The stimulus and filler sentences were randomly assigned to two stimulus lists with the constraints that two sentences

belonging to the same set were never in the same list, and each list contained an equal number of neutral and constraining sentences.

### Pictures

The pictures were line drawings from a normed database by Severens, Van Lommel, Ratinckx and Hartsuiker (2005). Each target picture was included as unrelated picture in another stimulus set. This way, we ensured that target pictures did not inherently draw more overt visual attention than unrelated pictures. The names of the objects in each display were never semantically associated with the verb in the neutral condition and only the target object could be associated with the verb in the constraining condition (association norms from Deyne, Navarro & Storms, 2013). The onsets of the names of objects in one display were never identical, nor were they identical to the onsets of the accompanying verbs.

Three repeated-measures ANOVAs with language (native, non-native) and picture type (target, distractor) as factors showed that object names were matched for frequency, phoneme count, and syllable count across languages and conditions ($ps > .10$) (Table 2). The selected object names were orthographically dissimilar (normalized orthographic Levenshtein distance $\leq .50$, $M = .15$, $SD = .13$[3]). The pictures had a mean H-statistic (a name agreement index) in Dutch of .62 ($SD = .49$)

---

[2] To check whether the semantic association strength between the verb and the target picture name was stronger in the constraining than in the neutral condition, and whether the association strength was similar across languages, we obtained a measure for semantic association from the snaut tool (Mandera, Keuleers, & Brysbaert, in press). In snaut, the association strength between verb and target is calculated based on co-occurrences in large text corpora. The stronger the association strength, the lower the measure.

As expected, paired t-tests pointed out that there was a stronger semantic association between the verbs and targets in the constraining condition than in the neutral condition ($p < .001$ for Dutch and $p = .002$ for English). Also, there was no significant difference between the association strengths in our English and Dutch stimuli ($p = .18$), indicating that our stimuli sentences were matched for semantic association strength between languages.

[3] 0 = no overlap, 1 = identical (Schepens, Dijkstra, Grootjen, & van Heuven, 2013).

Table 2. *Mean lexical characteristics of Dutch (native) and English (non-native) stimuli*

|  |  | Frequency[a] | Phoneme Count[b] | Syllable Count[b] |
|---|---|---|---|---|
| Picture name Dutch | Target | 4.39 (.61) | 4.17 (1.54) | 1.44 (.71) |
|  | Distractor | 4.28 (.29) | 4.19 (.76) | 1.31 (.37) |
| Picture name English | Target | 4.46 (.58) | 4.17 (1.20) | 1.44 (.62) |
|  | Distractor | 4.29 (.27) | 4.15 (.68) | 1.48 (.26) |
| Verb Dutch | Neutral | 3.85 (.60) | 5.28 (1.56) | 1.44 (.62) |
|  | Constraining | 3.48 (.77) | 4.83 (1.04) | 1.44 (.51) |
| Verb English | Neutral | 3.74 (.42) | 4.78 (.94) | 1.22 (.43) |
|  | Constraining | 3.50 (.62) | 4.78 (.88) | 1.33 (.49) |

*Note.* Standard deviations are indicated in parentheses.
[a]Zipf value (log10(frequency per million∗1000) (van Heuven, Mandera, Keuleers, & Brysbaert, 2014) retrieved from the SUBTLEX-US and SUBTLEX-NL databases (Brysbaert & New, 2009; Keuleers, Brysbaert, & New, 2010)
[b]CELEX database (Baayen, Piepenbrock, & Gulikers, 1995).

(Severens et al., 2005)[4]. To our knowledge, no name agreement scores are available for the picture set for bilinguals in L2.

### Sentences

Simple four-word SVO sentences were constructed for this experiment. The subject of the sentence was kept constant across all trials (*Mary* in English, *Marie* in Dutch). Repeated measures ANOVAs with language (native vs. non-native) and condition (neutral vs. constraining) as factors showed that verb frequency, phoneme count, and syllable count were matched across languages and conditions (all $ps \geq .10$). Table 2 reports the lexical characteristics of the stimuli in English and in Dutch.

The article preceding the sentence final noun was always indefinite, and English nouns never started with a vowel. This ensured that the article could not be used as a prediction cue.

### Recordings

Sentences were recorded in a sound attenuating room. A female native speaker of Dutch (34 years old) who majored in Dutch and English linguistics and literature at university pronounced the sentences for both the English and the Dutch recordings. English monolinguals rated her accent as 5.3 on a scale from 1 (very foreign accent) to 7 (native accent). We chose this speaker for our study because of her clear pronunciation in Dutch and English, and experience in recording psycholinguistic stimuli. Each sentence was recorded three times; the recording that we judged to have the most neutral prosody was selected for the experiment.

[4] The mean H-statistic of the full picture set of Severens et al. (2005) was 1.00 with scores ranging from 0 to 3.19. Lower H-statistic scores indicated higher name agreement.

The length of the recording frames starting at verb offset, and ending at noun onset initially differed significantly between Dutch and English ($t(35) = 10.87$, $p < .001$). In the non-native condition, participants would therefore have less time to generate predictions about upcoming referents than in the native condition. To eliminate this confound, the fragment was lengthened by a factor 1.2 for the English sentences and shortened by a factor 0.8 for the Dutch sentences, using Praat (Broersma & Weenink, 2014). This way, the length of the recording fragments was matched across languages ($ps \geq .10$). The mean length of the verb onset – noun onset frame was now 691 ms in Dutch and 708 ms in English. None of the participants indicated having noticed the manipulation of the auditory stimuli.

### Procedure

Participants were seated at a comfortable distance from the screen. They received written and verbal instructions to listen carefully to the sentences and to look at whatever they wanted as long as their gaze would not leave the screen (Huettig & Altmann, 2005; McQueen & Huettig, 2012). There was no explicit task. Eye movements were recorded from the right eye with an Eyelink 1000 eye-tracker (SR Research) with a sampling frequency of 1000 Hz. After successful calibration, the experiment began with two practice trials.

A fixation cross appeared on the screen for 500 ms, followed by the presentation of the four pictures in a two-by-two grid on the screen. Picture location was randomized. The auditory stimulus started to play 2200 ms after picture onset. This time lag was included to ensure that participants had enough time to see every object on the screen before verb onset. The trial ended when the sentence finished, and the next trial was started
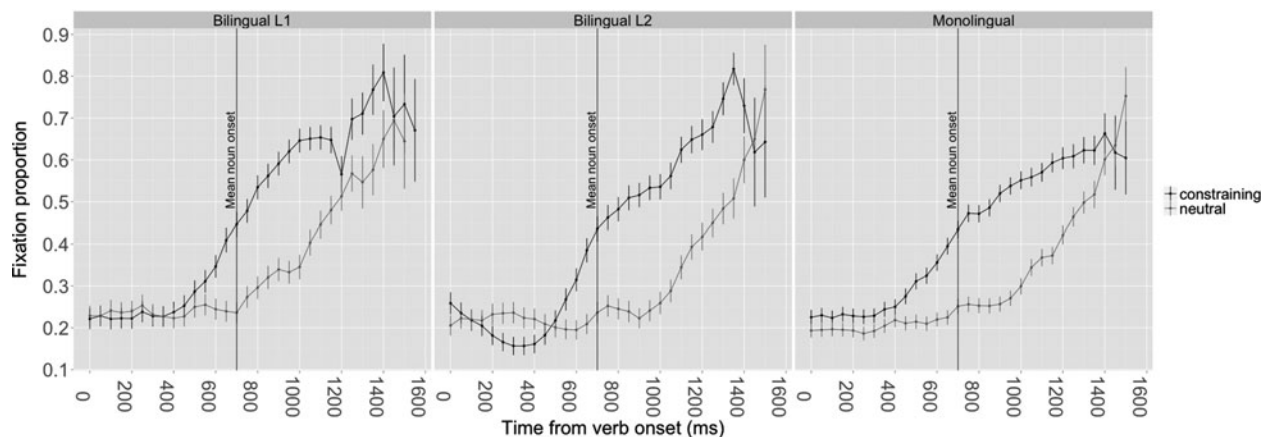
Figure 2  Results. Time course of fixation probability to target by condition (constraining, neutral) starting from verb onset for bilinguals in L1, bilinguals in L2, and for monolinguals. Note: Whiskers indicate the mean ± standard error.

by the experimenter after drift correction. Bilingual participants were presented with the stimuli in one of the lists in a Dutch (native) block and with the other list in an English (non-native) block. Language and list order were counterbalanced. Monolingual participants were presented with the stimuli of one list in the first block and with the stimuli of the other list in the second block. Both lists were presented to the monolinguals in English. List order was counterbalanced. In each block, the participants heard nine constraining and nine neutral sentences. Across the two blocks, none of the verbs were repeated, but the object displays were repeated. The eye tracker was recalibrated between the two blocks. The entire experiment took approximately 17 minutes.

After the experiment, participants completed the following additional tests: LexTALE Dutch, LexTALE English (Lemhöfer & Broersma, 2012) (see Table 1 for results), backward translation of the English verbs used in the experiment (bilinguals only), backward translation of the English nouns used in the experiment (bilinguals only), and the LEAP-Q language background questionnaire (Marian et al., 2007). The tests were presented in that order on a Macbook in a quiet room. Completion of the additional tests took approximately 25 minutes.

## Results

Figure 2 shows the time-course of target fixation as a function of condition for each language and speaker group. These probabilities reflect the number of samples of eye-data within a 50 ms time bin in which there was a fixation on the target picture, averaged over subjects and items.

The graph shows that participants were more likely to fixate on target objects in the constraining condition than in the neutral condition. Fixation proportions for the con-

straining and neutral conditions start to diverge well before the mean noun onset time in each of the three groups.

The starting point of the time frame for our analysis was chosen based on visual inspection of a plot of the time-course of the grand mean of fixation probability (over languages and listener types) and was defined as the first 50 ms time bin after verb onset in which the grand mean fixation probability began a rising trend (Barr, 2008). This method is conservative because by using the grand mean the choice cannot be biased by any hypothesis (Barr, 2008). As it takes approximately 200 ms to plan and execute a saccade (e.g., Matin, Shao & Boff, 1993; Saslow, 1967), we can assume that fixations that started earlier than 200 ms after noun onset were anticipatory in nature. Thus, the time frame for the analysis started at 350 ms after verb onset and ended 200 ms after noun onset. Each trial's individual verb onset and noun onset times were used to select the data. In addition to the analysis of the full time frame we analysed the data of the first four hundred milliseconds of data in the analysis frame aggregated into 100 ms time bins. This way, we tested when the effect of condition became significant in each group. In 3.39% of the samples in time frame from verb onset until 200 ms after noun onset there was a blink and 0.17% percent of the samples were out-of-screen. The out-of-screen and blink samples were included in the total sample count used to calculate proportions of looks to the target image.

The proportions of samples in the analysis time-frame in which there was a fixation to the target image were transformed using the empirical logit formula (Barr, 2008). Our data set was analyzed with linear mixed effects models with the lme4 (version 1.1-8), car (2.0-25) and lmerTest (version 2.0-25) package of R (3.2.1) (R Core Team, 2013). This allowed for inclusion of participant, sentence and target image as random factors (Baayen, Davidson & Bates, 2008).

For the analyses between languages in bilinguals, the fixed experimental factors were condition (constraining or neutral) and language (Dutch or English). The control variables list (A or B) and block (1 or 2) were also included as fixed factors. The models included random intercepts for participant, sentence and target picture. In each analysis we first fitted a model including all the fixed factors and interactions as well as the random intercepts for participant, sentence and target picture. If there was a significant effect of a factor, we added that factor as random slope for participant, sentence and target picture. For the comparison between listener types (monolinguals and bilinguals) in English and in Dutch listening, the fixed factors were condition (constraining or neutral) and listener type (monolingual or bilingual). All other factors were the same as in the within participants analysis[5]. To test whether there were any effects of English proficiency on predictive processing we compared each model without the factor lexTALE score (English) to the model with the factor lexTALE score and LexTALE as random slope for sentence and target picture using a likelihood ratio test. Eighteen trials were removed from the dataset because the verb was not translated correctly in the translation task that was performed after the main task, by that particular participant.

### *Comparison within Bilinguals (L1 vs. L2)*

The fixation proportion was significantly higher in the constraining condition than in the neutral condition ($\beta = -0.54, SE = .12, t = -4.49, p < .001$), confirming our prediction manipulation. There was no significant interaction between language (L1 vs. L2) and condition

(constraining vs. neutral) ($\beta = 0.04, SE = .10, t = .40, p = .69$). Nor were there any other significant main effects[6]. English proficiency (lexTALE) score did not significantly improve the model fit ($\chi^2(19) = 15.2, p = .71$)[7].

Separate analyses for each language revealed that the effect of condition was significant in L1 ($\beta = -0.65, SE = 0.17 \ t = -3.86, p = .001$)[8], and also in L2 ($\beta = -0.56, SE = 0.15, t = -3.58, p < .001$).

### *Comparison between L1 Monolingual listening (English) and L2 Bilingual listening (English)*

The fixation proportion was significantly higher in the constraining condition than in the neutral condition ($\beta = -0.69, SE = .12, t = -5.76, p < .001$). The effect of condition did not interact with listener type (monolingual versus bilingual) ($\beta = -.11, SE = .11, t = -.93, p = .36$). Nor were there any other significant main effects. English proficiency (lexTALE) did not significantly improve the model fit ($\chi^2(22) = 24.72, p = .32$). The effect of condition was also significant in the data of the monolinguals only ($\beta = -.79, SE = .16, t = -4.87, p < .001$).

### *Comparison between L1 Monolingual Listening (English) and L1 Bilingual Listening (Dutch)*

The fixation proportion was significantly higher in the constraining condition than in the neutral condition ($\beta = -.72, SE = .13, t = -5.57, p < .0001$). There was no significant interaction between listener type (monolingual vs. bilingual) and condition ($\beta = -.07, SE = .12, t = -.61, p = .55$). Proficiency (English LexTALE score) did not contribute significantly to the model fit ($\chi^2(22) = 29.21, p = .14$).

### *Time course analyses*

In the bilinguals, the effect of condition became significant in the third 100 ms time bin of the analysis time frame (550–650 ms) ($\beta = -.45, SE = .15, t =$

---

[5] In addition to our main analysis with the dichotomous factor 'condition' (neutral versus constraining), we checked whether there was an effect of the semantic association strength between verb and target pairs on fixation proportion to target images in the analysis time frame. The measure 'semantic association strength' was obtained from snaut (Mandera, Keuleers, & Brysbaert, in press) (see Footnote 2). We tested this for each analysis separately: within bilinguals (L1 and L2), between listener types (monolinguals and bilinguals) in English, and between listener types in L1 (Dutch for bilinguals and English for monolinguals).

In the within-bilinguals (bilinguals in L1 and L2) analysis, there was a marginally significant effect of association strength $\beta = -3.44, SE = 1.64, t = -2.09, p = .056$. The stronger the association strength, the more fixations to the target image in the analysis time frame. In the between listener type analysis (English in monolinguals and bilinguals), the main effect of association strength reached significance: $\beta = -3.37, SE = 1.49, t = -2.27, p = .032$. Finally, in the within L1 analysis (Dutch in bilinguals, English in monolinguals), no significant effect of association strength was found ($\beta = -2.03, SE = 1.38, t = -1.47, p = .14$). Importantly, there were no significant interactions between association strength and language or listener type in any of the analyses. The analyses suggest that stronger semantic association yields stronger prediction. We currently have no theory as to why the effect of association strength on target fixations did not reach significance in the within L1 analysis.

[6] After processing L2, processing in L1 tends to be slowed down (Misra, Guo, Bobb, & Kroll, 2012). Therefore, anticipatory effects were expected to be smaller or start later in an L1 block following an L2 block than vice versa. Our analyses showed that the order of language blocks did not interact with the effect of prediction, neither in the analysis of the entire time frame nor in the time course analyses.

[7] The result is reported for the model with LexTALE as random slope for target picture, but not for sentence. Condition was also included as random slope for participant and target picture. This was the maximum random effect structure justified by our sample (including LexTALE as random slope for sentence resulted in a model convergence error).

[8] The result is reported for the model with condition as random slope for participant, but not for target image. This was the maximum random effect structure justified by our sample.

−2.94, $p = .007$). There was no significant interaction between language and condition ($\beta = -.03$, $SE = .12$, $t = -0.21$, $p = .84$). In a separate analysis of the bilingual data for each language, the effect of condition also became significant in the third time bin of the analysis frame in English (550–650 ms after verb onset) ($\beta = -.47$, $SE = .20$, $t = -2.32$, $p = .03$) and in Dutch ($\beta = -.43$, $SE = .19$, $t = -2.24$, $p = .03$).

In the comparison between listener types in English (L1 monolinguals vs. L2 bilinguals) the main effect of condition was not yet significant in the first two time bins (350–450 ms after verb onset: $\beta = .06$, $SE = .13$, $t = .46$, $p = .65$, 450–550 ms after verb onset: $\beta = -.19$, $SE = .14$, $t = -1.35$, $p = .18$). However, the interaction between listener type and condition was significant in the first bin ($\beta = -.25$, $SE = .12$, $t = -2.09$, $p = .04$)[9], and marginally significant in the second bin ($\beta = -.22$, $SE = .12$, $t = -1.89$, $p = .06$). In the third time bin, the effect of condition became significant ($\beta = -.55$, $SE = .15$, $t = -3.78$, $p < .001$), and the interaction between listener type (monolingual vs. bilingual) and condition was no longer significant ($\beta = -.06$, $SE = .11$, $t = -.56$, $p = .57$).

Finally, we compared the two listener types in L1 (English in monolinguals vs. Dutch in bilinguals). The effect of condition became significant in the second time bin in the analysis frame $\beta = -.28$, $SE = .13$, $t = -2.09$, $p = .04$. The interaction between listener type and condition did not reach significance $\beta = -.13$, $SE = .12$, $t = -1.02$, $p = .31$.

In a separate analysis of the monolingual data, the effect of condition was significant for the first time in the second time bin in the analysis frame (450–550 ms after verb onset) ($\beta = -.41$, $SE = .18$, $t = -2.29$, $p = .03$). At that time, the effect was not yet significant for the bilinguals in English (L2) ($\beta = -.04$, $SE = .17$, $t = -.21$, $p = .83$) or in Dutch (L1) ($\beta = -.15$, $SE = .18$, $t = -.88$, $p = .38$).

## Discussion and conclusion

This study asked whether bilinguals predict information about upcoming referents on the basis of semantic context information during non-native comprehension, as monolinguals do in L1 comprehension. Following monolingual studies (e.g., Altmann & Kamide, 1999), we found that bilinguals use linguistic context information to generate predictions about upcoming referents in their non-native language (English). This effect was

of comparable magnitude in L1 listening in the same participants (Dutch) and in L1 listening by monolinguals in the same language (English). In addition, bilinguals listening in L1 (Dutch) predicted upcoming semantic information to a similar extent as monolinguals listening in L1 (English). English proficiency (lexTALE score) did not affect the prediction process. These findings confirm that bilinguals listening to non-native input are able to rapidly integrate auditory and visual input to constrain the subsequent domain of reference[10]. Consistent with the weaker links hypothesis (Gollan et al., 2008), time-course analyses suggested that bilinguals listening in either L1 or L2 predicted upcoming information slightly slower than monolinguals.

Kaan (2014) argued that predictive processing in a non-native language is not inherently different from predictive processing in the native language, but that other factors associated with non-native processing (e.g., cross-linguistic competition, inconsistent lexical representations in L2) can modulate prediction. No modulation of the prediction effect in the non-native language was found in the present study. Perhaps the modulating factors Kaan discussed only play a role under specific circumstances such as in sentences with infrequent words or cognates. Infrequent words are likely to have inconsistent representations because they are practiced less often. Also, no large cross-linguistic interference effects were expected because target words were never cognates. Furthermore, in visual world paradigm experiments like the present one, prediction processes may be facilitated as compared to EEG studies (e.g., Foucart et al., 2015), because visual candidates for prediction (pictures) are provided with each sentence (Kamide, 2008). Target words or target semantics were likely to be pre-activated along with the three other candidates.

Like us, Foucart et al. (2015) found a significant prediction effect in L2 speech processing using an EEG paradigm. The authors measured the modulation of the N400 effect elicited by an article that was gender congruent or incongruent with the predicted noun in L2 listening. The article-noun agreement rule manipulated in this experiment exists both in the bilingual participants' L1 and L2. Foucart et al. therefore suggested that prediction can be accomplished in L2 processing if the L2 is similar to the L1. Unlike in Foucart et al.'s study, no cognates were included as target words in our

---

[9] The result is reported for the model with condition as random slope for participant and target image. Listener type was included as random slope for sentence, but not for target image. This was the maximum random effect structure justified by our sample.

[10] Note that in the current design it is possible that sometimes the target picture (visible before the onset of the auditory signal) primed the verb, because of a strong semantic association between verb and target picture. This could strengthen the further prediction. It is impossible to dissociate the effect of associative strength between verb and target on verb priming vs. target prediction. As association strength did not differ between languages or listener groups (see Footnote 2), our conclusions still stand.

visual world experiment. Therefore, the prediction effect found in non-native listening in our experiment did not depend on target similarity between languages. However, English and Dutch are typologically similar languages, therefore Foucart et al.'s suggestion that prediction in L2 is facilitated by L1 and L2 similarity is still viable. The present experiment complements Foucart et al.'s results because we make a direct comparison between the prediction effect in bilinguals in L1, in L2 and in monolinguals in L1, and show that the magnitude of the prediction effect is the same in each language and speaker group.

Hopp (2015) also found an effect of prediction in non-native listening using a visual world paradigm. Unlike native listeners whose predictions were based on semantic and case-marking information, the non-native listeners were unable to use case-marking information to modulate predictions. Non-native listeners' predictive looks to likely patient objects may have been based on the semantic information extracted at the first NP in the sentence regardless of verb semantics, or on a combination of semantic information of the first NP and verb semantics. In the present experiment no picture of the first NP in the sentence was shown in the display, and only the verb distinguished the neutral from the constraining condition. Therefore, this study confirms that bilinguals listening in L2 can use verb semantics in order to predict features of upcoming input to the same extent in L1 and L2.

Previous studies showed that bilinguals have difficulty with predicting L2 input based on morphosyntactic information such as case or gender information (Dussias et al., 2013; Hopp, 2013, 2015). Predicting upcoming words together with morphosyntactic information (the gender of an article) is also difficult for bilinguals (Martin et al., 2013), unless the second language shares morphosyntactic features (e.g., gender-noun agreement rules) with the first (Foucart et al., 2014, 2015). However, in line with Koehne and Crocker (2015) the results of the present study show that bilinguals have no difficulty predicting input based on semantic information. This suggests that bilinguals predict to a similar extent in L2 as monolinguals do in L1, but that problems arise only when morphosyntax is involved, perhaps because of difficulty applying morphosyntactic agreement rules online quickly enough. An interesting question for future research would be whether increased processing speed (e.g., increased speech rate) would lead to difficulty using semantic information to generate predictions in L2 as well.

Speaker accent can affect speech processing (Adank, Evans, Stuart-Smith & Scott, 2009; Lagrou et al., 2013b; Weber, Betta & McQueen, 2014). Dutch–English bilinguals in Belgium are frequently exposed to non-native speakers in school and work settings, and in the media. Therefore, they are familiar with Dutch-accented English like the accent of the speaker in the experiment. A previous study from our lab (Lagrou et al., 2013b) showed that, in a lexical decision task, Dutch–English bilinguals responded faster to English stimuli pronounced by a native speaker than to English stimuli pronounced by a non-native speaker. If words are recognized more slowly by L2 listeners when pronounced by an L2 speaker than by an L1 speaker (Lagrou et al., 2013b), then an interaction effect of language (L1 or L2) with prediction of upcoming information is likely to be more pronounced when the speaker of the experimental stimuli is a non-native speaker. No such interaction was found in the present experiment. Whether various strengths of non-native accents affect the prediction process differently in L1 and L2 listeners remains an open issue[11].

English proficiency as measured with LexTALE (Lemhöfer & Broersma, 2012) did not affect the magnitude of the prediction effect in bilinguals and monolinguals. This may be due to the high level of proficiency of our participants, although these were still clearly unbalanced bilinguals who only use L2 during a small proportion of their time. Alternatively, there may not have been sufficient variance (bilinguals $M = 78.5$, $SD = 10.49$) to detect an interaction effect of proficiency with prediction skill. Conversely, production skill and not recognition skill may be an indicator of prediction skill (Mani & Huettig, 2012), and the LexTALE does not tap into production skill directly. In any case, the present data show that the proficiency level of these unbalanced bilinguals suffices for predictive language processing similar to that in the native language.

The time course analyses showed that prediction effects reached significance 100 ms later for bilinguals (in both languages) than for monolinguals. One theoretically interesting interpretation would be that activation and

---

[11] An interesting way to assess whether prediction was affected by speaker accent, is to look at the prediction effect throughout the course of the English part of the experiment. If there was an effect of accent, listeners may have adapted to the speaker accent throughout the English block (although other factors such familiarity with the task and experimental design may also yield such adaptations).

We checked whether there was an effect of the time course of the experiment by testing whether the effect of condition was larger in the second half of the English block than in the first half of the English block, both for bilinguals and monolinguals. The factor 'experiment half' (first half vs. second half) was added to the model used for the analysis of the English (monolingual and bilingual) data. The interaction effect between experiment half and condition (constraining vs. neutral) was not significant ($\beta = .03$, $SE = .21$, $t = -.16$, $p = .87$). The main effect of experiment half was also not significant ($\beta = -.41$, $SE = .22$, $t = -1.90$, $p = .07$). No difference was found between the proportion of fixations on the target image or predictive behavior in the first and second half of the English block.

prediction develops slower for bilinguals. However, this may also merely be due to lack of power in the bilingual data sets. The monolinguals were exposed to both stimuli lists in English whereas the bilinguals were exposed to one list in each language. Therefore, the monolingual data set is twice the size of the bilingual data sets of each language, which increases power to detect effects. However, the delay of one time bin also exists in the full bilingual data set (English and Dutch combined), which is equal in size to the monolingual data set. This supports that there may not just be a power issue, but that in fact bilinguals predicted upcoming information slightly less rapidly than monolinguals. This would be consistent with the weaker links hypothesis of bilingual language processing, which states that division of use between a bilingual's two languages results in weaker links between lexical items' semantics and phonology (Gollan et al.,

2008). This should result in slower lexical access and could possibly lead to slower predictions during language comprehension.

This study shows that L2 listeners use semantic information provided by sentences to restrict the expected subsequent domain of reference to the same extent as in L1 processing by bilinguals and monolinguals. This finding suggests that, when no grammatical rules need to be processed online in order for participants to generate a prediction, the basic principles of recent theories of prediction in language comprehension (cf. Altmann & Mirković, 2009; Federmeier, 2007; Kutas et al., 2011; Pickering & Garrod, 2013) also apply to L2 processing in highly proficient bilinguals. Future studies will have to point out more precisely in what circumstances predictive language processing is retained in L2 processing, and when it is not.

Appendix A. *Stimuli sets in English and Dutch.*

| Set | Constraining English (Dutch) | Neutral English (Dutch) | Target object English (Dutch) | Distractor 1 English (Dutch) | Distractor 2 English (Dutch) | Distractor 3 English (Dutch) |
|---|---|---|---|---|---|---|
| 1 | cuts (snijdt) | shares (deelt) | cheese (kaas) | mailbox (brievenbus) | desk (bureau) | bottle (fles) |
| 2 | reads (leest) | steals (pikt) | letter (brief) | car (auto) | backpack (rugzak) | wheelchair (rolstoel) |
| 3 | tastes (proeft) | grabs (grijpt) | carrot (wortel) | dress (jurk) | napkin (zakdoek) | ruler (lat) |
| 4 | climbs (beklimt) | fixes (repareert) | wall (muur) | lock (slot) | heel (hak) | shower (douche) |
| 5 | boils (kookt) | buys (koopt) | potato (aardappel) | coat (jas) | flower (bloem) | dog (hond) |
| 6 | visits (bezoekt) | prefers (verkiest) | church (kerk) | skirt (rok) | sausage (worst) | belt (riem) |
| 7 | closes (sluit) | reveals (onthult) | window (raam) | horse (paard) | needle (naald) | spoon (lepel) |
| 8 | trains (traint) | catches (vangt) | horse (paard) | letter (brief) | plate (bord) | sword (zwaard) |
| 9 | folds (vouwt) | saves (bewaart) | dress (jurk) | bucket (emmer) | candle (kaars) | chest (kist) |
| 10 | knits (breit) | loses (verliest) | scarf (sjaal) | cheese (kaas) | comb (kam) | barrel (ton) |
| 11 | sharpens (slijpt) | measures (meet) | pencil (potlood) | church (kerk) | rope (touw) | tree (boom) |
| 12 | builds (bouwt) | breaks (breekt) | fence (hek) | pencil (potlood) | necklace (ketting) | mirror (spiegel) |
| 13 | empties (leegt) | paints (verft) | mailbox (brievenbus) | window (raam) | roof (dak) | chair (stoel) |
| 14 | irons (strijkt) | designs (ontwerpt) | skirt (rok) | wall (muur) | bicycle (fiets) | kite (vlieger) |
| 15 | drives (bestuurt) | takes (neemt) | car (auto) | potato (aardappel) | knife (mes) | shovel (schop) |
| 16 | wears (draagt) | chooses (kiest) | coat (jas) | fence (hek) | mountain (berg) | bridge (brug) |
| 17 | opens (opent) | draws (tekent) | lock (slot) | carrot (wortel) | wig (pruik) | raft (vlot) |
| 18 | fills (vult) | throws (gooit) | bucket (emmer) | scarf (sjaal) | key (sleutel) | whistle (fluit) |

# References

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 520.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73,* 247–264.

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science, 33,* 583–609.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database. Release 2 (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11,* 280–289.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59,* 457–474.

Boland, J. E. (2005). Visual arguments. *Cognition, 95,* 237–274.

Broersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.78). Retrieved from www.praat.org

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977–990.

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,* 1029–1040.

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. Groot & J. F. Kroll (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 279–300). Lawrence Erlbaum.

Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369,* 20120394–20120394.

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

Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods, 45,* 480–498.

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). WHEN GENDER AND LOOKING GO HAND IN HAND. *Studies in Second Language Acquisition, 35,* 353–387.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44,* 491–505

FitzPatrick, I., & Indefrey, P. (2007). Effects of sentence context in L2 natural speech comprehension. *Proc. of the Cognitive Neuroscience Master of the Radbout University, 2,* 43–56.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1461–1469.

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition,* 1–7.

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,* 787–814.

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

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

Hopp, H. (2015). Semantics and Morphosyntax in Predictive L2 Sentence Processing. *International Review of Applied Linguistics*.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition, 96,* 23–32.

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

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? In J. Rothman & S. Unsworth (eds.), *Linguistic Approaches to Bilingualism* (pp. 257–282).

Kamide, Y. (2008). Anticipatory Processes in Sentence Processing. *Language and Linguistics Compass, 2,* 647–670.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42,* 643–650.

Koehne, J., & Crocker, M. W. (2015). The Interplay of Cross-Situational Word Learning and Sentence-Level Constraints. *Cognitive Science, 39,* 849–889.

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).

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013a). Interlingual lexical competition in a spoken sentence context: Evidence from the visual world paradigm. *Psychonomic Bulletin & Review, 20,* 963–972.

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013b). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition, 16,* 508–517.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44,* 325–343.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106,* 1126–1177.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4.*

Mandera, P., Keuleers, E., & Brysbaert, M. (in press). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language.*

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 843–847.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50,* 940–967.

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

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics, 53,* 372–380.

McQueen, J. M., & Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognized. *The Journal of the Acoustical Society of America, 131,* 509–517.

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language, 67,* 224–237.

Mitsugi, S., & Macwhinney, B. (2015). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition, 19,* 19–35.

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

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics, 77,* 720–730.

Saslow, M. G. (1967). Latency for saccadic eye movement. *JOSA, 57*(*8*), 1030–1033.

Schepens, J., Dijkstra, T., Grootjen, F., & van Heuven, W. J. B. (2013). Cross-Language Distributions of High Frequency and Phonetically Similar Cognates. *PLoS ONE, 8,* e63006.

Severens, E., Lommel, S. V., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica, 119,* 159–187.

Van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad – any implications for neuropragmatics. *Italian Journal of Linguistics, 22,* 181–208.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67,* 1176–1190.

Weber, A., Betta, A. M. D., & McQueen, J. M. (2014). Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners. *Journal of Phonetics, 46,* 34–51.

Weber, A., & Broersma, M. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (ed.), *The Encyclopedia of Applied Linguistics.* Hoboken, NJ, USA: John Wiley & Sons, Inc.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50,* 1–25.

Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A., & Duyck, W. (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1579–1586.

Zirnstein, M., Van Hell, J. G., & Kroll, J. F. (2015, May). *How Bilingualism and Cognitive Control Impact L2 Reading Comprehension: Evidence from ERPs.* Poster presented at the International Symposium on Bilingualism.