Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–14

DOI: 10.1177/1747021820948019
qjep.sagepub.com
SAGE

Effects of speech rate on anticipatory eye movements in the visual world paradigm: Evidence from aging, native, and non-native language processing

Leigh B Fernandez[1], Paul E Engelhardt[2], Angela G Patarroyo[3] and Shanley EM Allen[1]

Abstract

Research has shown that suprasegmental cues in conjunction with visual context can lead to anticipatory (or predictive) eye movements. However, the impact of speech rate on anticipatory eye movements has received little empirical attention. The purpose of the current study was twofold. From a methodological perspective, we tested the impact of speech rate on anticipatory eye movements by systemically varying speech rate (3.5, 4.5, 5.5, and 6.0 syllables per second) in the processing of filler-gap dependencies. From a theoretical perspective, we examined two groups thought to show fewer anticipatory eye movements, and thus likely to be more impacted by speech rate. Experiment 1 compared anticipatory eye movements across the lifespan with younger (18–24 years old) and older adults (40–75 years old). Experiment 2 compared L1 speakers of English and L2 speakers of English with an L1 of German. Results showed that all groups made anticipatory eye movements. However, L2 speakers only made anticipatory eye movements at 3.5 syllables per second, older adults at 3.5 and 4.5 syllables per second, and younger adults at speech rates up to 5.5 syllables per second. At the fastest speech rate, all groups showed a marked decrease in anticipatory eye movements. This work highlights (1) the importance of speech rate on anticipatory eye movements, and (2) group-level performance differences in filler-gap prediction.

Keywords

visual world paradigm; anticipatory eye movements; wh-movement; bilingualism; aging; speech rate

Received: 30 September 2019; revised: 18 June 2020; accepted: 28 June 2020

In the current study, we had a (primary) methodological research objective and a (secondary) theoretical research objective. Our methodological objective involves speech rate, which is not consistently reported in visual world paradigm (VWP) studies. Therefore, the existing literature contains a confound, which means that comparisons across studies are problematic. We tested the impact of speech rate on anticipatory eye movements by systemically varying speech rate, within the "normal" range, from 3.5 to 6.0 syllables per second. Our secondary objective was theoretical and involves groups that are likely to be differentially impacted by speech rate: older adults and L2 speakers. Older adults are affected by age-related declines in various language abilities, and L2 speakers have less proficiency in L2. To investigate anticipatory behaviour in these two groups, we examined the processing of filler-gap dependencies. Because this study has two research objectives, and because it touches on a variety of different topics, we feel it is important to clearly outline the structure of the Introduction leading up to the experiments. With respect to the methodological objective, it is necessary to introduce the VWP and the few studies that have at least considered the role of speech rate in anticipatory language processing. We then turn to the specific syntactic constructions that were used to drive anticipatory eye movements (i.e., filler-gap dependencies). Finally, with respect to the theoretical

[1]Technische Universität Kaiserslautern, Kaiserslautern, Germany
[2]University of East Anglia, Norwich, UK
[3]Humboldt-Universität zu Berlin, Berlin, Germany

Corresponding author:
Leigh B Fernandez, Technische Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany.
Email: leigh.fernandez@sowi.uni-kl.de

objective, we review the literature on filler-gap dependencies in native and non-native speakers and provide an overview of the mechanisms underlying age-related declines in language processing.

Language processing in the visual world

In the VWP, participants' eyes are tracked as they listen to utterances while viewing an array of objects. Language processing typically requires the integration of multiple sources of information to generate grammatically accurate representations of incoming linguistic input (Tanenhaus et al., 1995). Studies have shown that the parser not only integrates each word into the preceding structure as it is encountered, but also actively makes predictions on the basis of preceding information (for reviews, see Ferreira & Chantavarin, 2018; Huettig, 2015; Huettig & Mani, 2016; Kamide, 2008; Kuperberg & Jaeger, 2016; Staub, 2015). Anticipatory behaviours in sentence processing are observed with millisecond-level accuracy by looking, for example, at eye movements to an object before it is explicitly referred to in the input, as is the case in the VWP. The assumption is that, as visual attention shifts to an object in the scene, it provides insights into real-time language comprehension (e.g., Altmann & Kamide, 1999; Tanenhaus et al., 1995).

In an early study, Allopenna et al. (1998) showed that phonological overlap created competition in looks to objects whose names shared the same initial onsets (e.g., *candy* and *candle*). Native speakers looked equally towards both objects until later phonological information disambiguated the reference. This suggests that listeners direct eye movements to potential referents of linguistic input as they process the acoustic signal (see also, Magnuson et al., 2007). This type of effect has been shown across a wide range of different variables (e.g., Altmann & Kamide, 1999, 2004; Altmann & Mirkovic, 2009; Huettig et al., 2011; Kamide, 2008; Knoeferle et al., 2005; Sedivy et al., 1999). Thus, the VWP provides temporally sensitive and incremental insights into anticipatory language comprehension.

Spoken language comprehension—speech rate

Because the VWP focuses on processing of spoken rather than written stimuli, it also allows insight into how suprasegmental features of speech such as speech rate, pitch, and stress can facilitate (or hinder) interpretations. Several studies have investigated properties of the speech signal using the VWP, finding that certain suprasegmental cues can be used with visual context to generate anticipatory eye movements (for a review, see Huettig et al., 2011). However, only one study has directly investigated the impact of speech rate, arguably the most salient suprasegmental feature (i.e., Huettig & Guerra, 2019). In a sampling of 45 VWP studies published in high-impact journals from 2008 to 2018,[1] only three reported the speech rate of their materials. These studies reported speech rates of 1.3, 2.6, and 3.1 syllables per second. Ten of the 45 studies noted that the stimuli were within a "normal" range but did not provide quantitative estimates, 16 studies reported some information on their utterances but reported nothing about speech rate (e.g., stimuli were recorded by a native speaker), and the remaining 16 did not include any information about their stimuli.

In one of the few studies to actually manipulate speech rate, Huettig and Guerra (2019) had Dutch speakers listen to sentences while looking at an array of four objects, and were asked to "Look at the . . ." The display contained three objects with the same gender and one of contrasting gender. When the contrasting object was the target, participants could predict the object at the gender-marked article (i.e., before the noun was mentioned). Participants anticipated the object when they heard the instruction at "slow" and "normal" speech rates when they saw the array for 4 seconds before hearing the instruction. When they previewed the array for 1 second, prediction was only observed in the "slow" condition. This confirms an important role of speech rate on prediction in the VWP.

These findings are in line with Ferreira and Chantavarin (2018), who argued that prediction effects should not be explained from atypical experimental settings, ". . . an experiment in which words are presented at unusually slow rates . . ." (p. 445). However, Huettig and Guerra did not provide speech rate (they reported the means of the duration of their two speech rates) and coupled with the lack of consensus on what "normal" speech rate is, it is difficult to compare their study with previous research (which tends to provide speed in syllables per second). To summarise, there is currently a major methodological limitation in the literature concerning speech rates in the VWP such that, at present, it is unknown how much speech rate impacts anticipatory eye movements. Therefore, in the current study we plan to directly test anticipatory eye movements across four different auditory presentation speeds. However, unlike Huettig and Guerra (2019), we did not vary preview time.

A wide range of "normal" speech rates have been reported in the literature. Several studies have reported that the average speech rate is between 2.50 and 4.56 syllables per second (e.g., Dickey et al., 2007; Fónagy & Magdics, 1960; Wilshire, 1999). Hertrich et al. (2013) reported that the normal speech rate is between 4.0 and 8.0 syllables per second, while Munro and Derwing (2001) reported that the optimal rate for comprehensibility (judged by L1 English speakers) was 4.76 syllables per second. Speech rate is impacted by several factors such as age, dialect, or whether you are speaking in your native or non-native language.

Likewise, comprehension is impacted by speech rate. Wingfield et al. (1985) investigated comprehension in older adults (65–73 year olds) vs. university students (18–22 year olds), finding that older adults showed a steeper decline in accuracy (in sentence recall) as speech rate increased. In L2 comprehenders, auditory input rate is the most often identified source of comprehension difficulty (Graham, 2006).

Filler-gap dependencies in native (L1) speakers

Filler-gap dependences such as wh-questions (e.g., Who$_i$ did the bride tickle t_i at the mall?) are formed by moving the filler (*who$_i$*) from its canonical position (e.g., The bride ticked *who* in the mall). When the filler is moved it leaves a gap (or trace) in its place (t_i), and the filler must be associated with the gap to be assigned a thematic role at the verb (i.e., who is doing what to whom).[2] The correct interpretation of a filler-gap dependency can only be formed after the gap is located. Research suggests that the parser actively searches for the gap as the sentence unfolds (e.g., Frazier, 1987; Frazier & Flores D'Arcais, 1989). Filler-gap dependencies therefore require top-down syntactic knowledge about where a filler can be posited, rather than bottom-up information that signals where to posit a gap (Atkinson et al., 2018).

Using the VWP to investigate filler-gap dependencies, Sussman and Sedivy (2003) investigated the processing of wh-movement by comparing fixations during yes/no sentences (1) and wh-questions (2) after participants listened to a story (about a girl squashing a spider with her shoe).

(1) Did Jody squash the spider with her shoe?
(2) What$_i$ did Jody squash *t_i the spider with t_i?

In (2) there is a potentially ambiguous gap location (*t_i) following *squash* due to the question (i.e., *What did Jody squash* . . .). However, this interpretation is ultimately rejected given that the utterance continues with *the spider*. Sussman and Sedivy found that participants actively tried to form a filler-gap dependency. This was based on the increased fixations on *the spider* around the onset of *squash*, suggesting an association between the filler with the first, but ultimately, incorrect gap. Later in the sentence, participants showed increased fixations to the correct answer (*the shoe*) around the onset of the preposition (*with*). Critically, this pattern of eye movements was not observed in the yes/no question that did not contain a filler-gap dependency.

Another type of filler-gap construction—the intermediate-gap construction—has been used to test L1 and L2 filler-gap dependency processing. In an intermediate-gap construction (3), the filler (*who$_i$*) and its gap site (t_i) span more than one clause, and these clauses are broken up by an intermediate gap (t'_i). In self-paced reading, Gibson and Warren (2004) found that L1 English speakers slowed down at the intermediate gap in (3), but showed shorter reading time at the final gap (following *pleased*) in (3) compared with (4). This suggests that the parser reactivated the filler at the intermediate gap site, which facilitated forming of the dependency at the final gap site. Gibson and Warren (2004) argued that intermediate gaps in constructions serve to break up the filler-gap dependency into two shorter dependencies, easing working memory load and facilitating the integration of the filler and the gap compared with a sentence like (4) with no intermediate gaps.

(3) The manager who$_i$ the consultant claimed t'_i that the new proposal had pleased t_i will hire five workers tomorrow.
(4) The manager who$_i$ the consultant's claim about the new proposal had pleased t_i will hire five workers tomorrow.

Language processing in non-native (L2) speakers

As reviewed above, research has found that individuals employ an active-filler strategy and rapidly build expectations during sentence processing. However, the majority of this research comes from native monolingual speakers. Given that half, if not more, of the world is bilingual (Grosjean, 2010), it is important that psycholinguistic theories account for bilingual processing. Two lines of L2 research are relevant. The first is filler-gap dependencies and the second is expectation building or prediction.

Several studies of L2 sentence processing have used filler-gap dependencies to test whether L2 speakers are capable of processing them in the same way as native speakers. Using self-paced reading and intermediate-gap constructions, Marinis et al. (2005) found that L2 speakers of English (L1s: Greek, German, Chinese, or Japanese) showed no facilitation from the intermediate gap. They concluded that L2 speakers do not use top-down syntactic knowledge the way L1 speakers do, but instead use a lexically driven bottom-up strategy to form a dependency between a filler and its lexical sub-categorizer. Later, Clahsen and Felser (2006) argued for a Shallow Structure Hypothesis, which suggests that L2 speakers rely less on syntactic information and, instead, rely more on lexical-semantic and pragmatic information. They assumed that at least two processing routes operate in parallel; one is a shallow heuristics-driven route while the other creates "deep/full" grammatical representations. The route that is employed may depend on several factors, such as proficiency or grammatical knowledge. Thus, L2 speakers may rely on the heuristic pathway more often than L1 speakers.

In contrast, other research has found that L2 speakers may be able to use the "grammatical route" and form syntactically driven filler-gap dependencies in some situations.

Dekydtspotter et al. (2006) re-analysed the data from Marinis et al. (2005), looking at the segment following the intermediate gap. They found evidence for delayed activation in two of the four L2 groups. Furthermore, Hopp (2006) argued that syntactic processing is not "shallow" for all L2 speakers, but rather dictated by L2 proficiency (which was not reported in Clahsen & Felser, 2006, or Marinis et al., 2005). Along the same line, Pliatsikas and Marinis (2013) found intermediate gap facilitation in a group of L2 speakers that had naturalistic exposure to English (i.e., lived in an English speaking country). Again, suggesting that L2 speakers can show native-like syntactic processing. Fernandez et al. (2018) used auditory presentation (arguably a more ecologically valid method than self-paced reading) to test intermediate-gap constructions. They found that L2 speakers showed similar facilitation as L1 speakers at the intermediate gap, which suggests the importance of suprasegmental information in language processing and according to Fernandez et al. resulted in "deeper" syntactic processing (i.e., more reliance on the grammatical route).

A second area of research relevant for the current study focuses on L2 speakers' ability to anticipate or build expectations in sentence processing. Some have argued that L2 speakers may not build expectations to the same extent as L1 speakers (e.g., Grüter et al., 2012, 2014, 2017). However, others have found that highly proficient L2s are capable of predicting upcoming information (e.g., Dussias et al., 2013; Hopp, 2013). In contrast, Kaan (2014) concluded that L1 and L2 speakers do not qualitatively differ in their expectation-building mechanism, but rather group differences derive from individual differences that affect language processing (e.g., processing strategies, quality of lexical representation, competing lexical information, etc.). Thus, expectation building differences may be linked to L2 proficiency or possibly some other individual difference(s) variable.

Sentence processing in older adults

Research suggests that certain cognitive capabilities decline with age, including inhibition, memory, processing speed, and reasoning (see, for example, Burke & Mackay, 1997; Rosselli et al., 2014; Salthouse, 2010; Shafto & Tyler, 2014). These factors—especially processing speed and working memory—have implications for language processing. However, other abilities remain static or increase with age (e.g., vocabulary and real-world knowledge), which may augment language processing (Stine-Morrow et al., 2006).

Eye movement patterns also change with age. Schik et al. (2000) investigated eye movements to visual targets in 90 adults ranging from 17 to 75 years of age, finding that the latency to initiate an eye movement to a target increased with age. In a reading study, Rayner et al. (2006) found that older adults (>70 years) read more slowly and made longer fixations than younger adults (<35 years). In addition, they found that older adults employed a "riskier" reading strategy in which they were more likely to skip words and make regressions, which they argued was due to building expectations of upcoming words based on contextual information (cf. Warrington et al., 2018). Consistent with this, in studies using auditory stimuli, it was found that older adults are more efficient at using context to compensate for degraded speech relative to younger adults (for review, see Pichora-Fuller, 2008).

Given that filler-gap dependencies require cognitive resources and working memory, older adults may process these structures more slowly than younger adults. It has been found that older adults are able to form filler-gap dependencies when working memory is controlled for (i.e., when the distance between the filler and gap were no further than five words apart; Zurif et al., 1995). However, the majority of research investigating filler-gap dependencies in older adults comes from research with acquired language disorders (where healthy older adults serve as control participants). For example, Dickey et al. (2007) investigated filler-gap dependencies (based on the Sussman and Sedivy materials) in individuals who had an acquired language disorder, and compared them to age-matched controls. They observed similar results in their controls as the young adults in Sussman and Sedivy. Thus, neither study directly compared younger and older (neuro-typical) adults, but both studies did show evidence for gap filling, and more-or-less similar patterns across age.

Current study

In this study, we had a (primary) methodological and (secondary) theoretical research objective. Methodologically, we tested the impact of speech rate on anticipatory eye movements. Previous research investigating speech rate has reported average speech rates between 2.5 and 8.0 syllables per second (e.g., Dickey et al., 2007; Fónagy & Magdics, 1960; Hertrich et al., 2013; Wilshire, 1999), and in the current study we varied speech rate from 3.5 to 6.0 syllables per second. Given that prior VWP research did not discuss or even report speech rate, we hope that this study will provide more scientifically supported and appropriate speech rate information for future studies. The lack of consistency in speech rate across studies may bias towards finding anticipatory eye movements with atypically slow speech or to not find anticipatory eye movements with atypically fast speech (Ferreira & Chantavarin, 2018). Theoretically, we compared anticipatory eye movements across groups of speakers that are likely to differentially impacted, for different reasons, by speech rate in the processing of filler-gap dependencies (Munro & Derwing, 2001; Wingfield et al., 1985). Experiment 1 tests whether speech rate affects anticipatory eye movements in the processing of filler-gap dependencies with native speakers of English ranging in age from 18 to 75 years of age, and Experiment 2 compares younger L1 speakers of English with late L2 speakers of English.

Experiment 1

Experiment 1 compared two groups of native speakers of English to test anticipatory eye movements across the lifespan, younger adults ranging from 18 to 24 years and older adults from 40 to 75 years. In this experiment, from a methodological perspective, we hypothesised that as speech rate increases, anticipatory eye movements to the target at the gap site will decrease. From a theoretical perspective, and based on prior research, we expect older adults will show a more substantial decrease in anticipatory eye movements as speech rate increases compared with younger adults. However, it could also be that older adults are able to use context more efficiently and commit to a "riskier" processing strategy, leading to a less dramatic decrease (or even similar) anticipatory eye movements compared with younger adults.

Method

Participants. Fifty-six native (L1) English speakers were recruited from Northumbria University and the greater Newcastle upon Tyne area. All participants had spoken English from birth and had not started to learn a second language before the age of 6. All had normal or corrected-to-normal vision. Thirty-one participants were in the older group with a mean age of 57.19 years ($SD = 11.17$, range 40–75 years) and 25 were in the younger group with a mean age of 19.32 ($SD = 1.35$, range 18–24 years). The older group had an average of 17.70 years of education ($SD = 4.28$), while the younger group had an average of 15.00 years of education ($SD = 0.89$). Given that years of education was not normally distributed for the younger group a Wilcoxon signed-rank test was run, this difference was significant $W = 572.50$, $p < .05$.

Materials

Linguistic stimuli. The experiment consisted of 53 trials: 3 practice trials, 20 critical trials, and 30 filler trials. Items were based on the Dickey and Thompson (2009) stimuli and consisted of a short story followed by one of two wh-movement comprehension probes (see Example 5).

(5) One day a wolf and a deer were sleeping near a cave. The wolf became crazed and the wolf attacked the deer. A hawk watched as the deer escaped.
   a. Who_i did the wolf attack t_i near the cave?
   b. Point to who_i the wolf was attacking t_i near the cave.

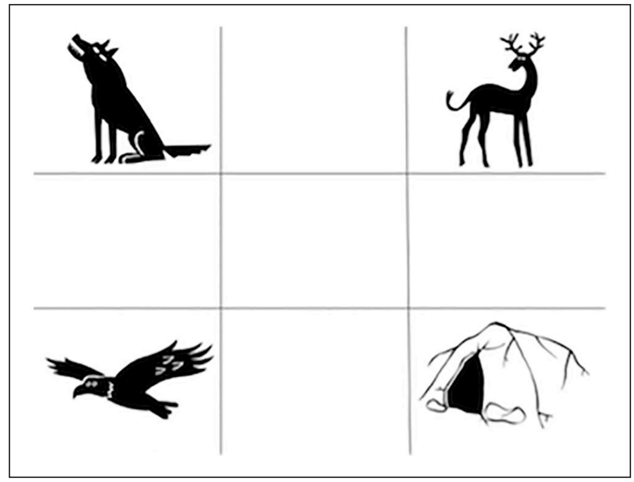

Figure 1. Example stimulus.

All stories were three sentences long. The first sentence introduced an event, the event location, and the agent and patient or theme of the event (which were animate NPs). The second sentence described the transitive event, and the third sentence contained an additional animate actor that served as a distractor. The animate NPs and location were matched for syllable length and did not contain initial phonological overlap. Two comprehension probe types were used to balance items that began with a wh-word or *point to* probe (see a and b in (5), (a) was an object-extracted wh-question and (b) was an object-extracted relative clause probe.

Items were constructed in a 2×4 design: probe type (point to/wh-word) and speed (3.5, 4.5, 5.5, and 6.0). Eight lists were created, such that items were rotated in a Latin square design. For the analysis, the two probe types were collapsed, resulting in a 1×4 design (see below for details). The fillers probed the subject (10 items), the distractor (10 items), and the location (10 items) mentioned in the story. Half of the filler probes began with a wh-word and half with *point to*. All items were recorded by a native speaker of British English with a Northeast accent at a sampling rate of 44,100 Hz. Speed was manipulated using Praat, making four conditions from the same item based on syllables per second.[3]

Visual stimuli. Fifty visual stimuli were created with four images depicting the three animate NPs and the location mentioned in the story (see Figure 1). The four images were black and white drawings obtained from Microsoft clipart. The image area was divided into nine equal-sized squares, and images were placed in the four corner squares positioned around a central fixation point. Images were rotated across items and were distributed across grid locations equally.

Apparatus. Stimulus presentation was programmed using Experiment Builder software, and eye movements were recorded with an Eyelink 1000 (SR Research Ltd., Mississauga, Canada) desktop mount sampling at 1,000 Hz. Viewing was binocular but only the right eye was tracked. Participants sat approximately 33 inches (83 cm) away from the screen, and a chinrest minimised head movements. Stimuli were presented on a 19 inch (48 cm) screen (60 Hz refresh rate) with an image size of 1024 × 769 pixels. Calibration was completed using a 9-point calibration/validation sequence.

Procedure. The experiment began with an information sheet and consent form. This was followed by a short language questionnaire, and then participants completed the eye-tracking study. Younger adults were given course credit and older adults were paid £5 for participating. Each participant was asked to listen to a story, while viewing an array of images that showed the objects mentioned in the story. (The images appeared at the same time the story began.) After listening to the story, participants had to answer a question about the story. Participants answered the question, after complete presentation of the question, by clicking on the image that they thought best answered the question only. The entire session lasted approximately 30–45 minutes.

Data analytic plan

Accuracy. Accuracy was recorded and coded as 1 (correct) or 0 (incorrect) and was submitted to a logit mixed effects model. Fixed effects for accuracy measures included group (younger/older) and speech rate (3.5, 4.5, 5.5, 6.0). Group was sum coded with 0.5/−0.5 and speech rate was sum difference coded such that each level was compared with the following level (3.5 vs. 4.5, 4.5 vs. 5.5, 5.5 vs. 6.0). The t and p values are provided in the results section, and additional model information is provided in the Supplementary Materials, Section 1a and all code and data are available online at https://osf.io/9s7ye.

Eye movements. To test anticipatory eye movements, we analysed eye movement data beginning 200 ms following the onset of the critical verb (*attack/attacking*), which accounts for the time it takes to programme an eye movement and extending outwards 1,000 ms. Evidence of active gap filling should occur at the onset of the verb in the form of more fixations to the theme (*the deer*) compared with the agent (*the wolf*) given where the parser should associate the filler with the gap (t_i). Due to the speech rate manipulation, the analysis window encompasses slightly different linguistic information.[4] However, given that the theme (nor the distractor) is never overtly mentioned, this time window captures the important parts of the (auditory) stimuli across all speeds. Note that the time between the onset of the verb and the onset of the following segment (*near the cave*, R in example 5) had a mean length of 380 ms (SD = 32). Table 1 shows the mean time between the onset of the verb and the following segment for each speech rate.

Table 1. Mean time (in milliseconds) between verb onset and the following segment.

Condition	Mean length in ms (*SD*)
3.5	523 (131)
4.5	382 (99)
5.5	318 (76)
6.0	295 (81)

SD: standard deviation.

Given that we were interested in the pattern of fixations across time, the empirical logit of fixation counts was used for visualisation and data analysis. The empirical logit is the log-odds ratio of looking at the target (compared with looking not at the target, that is, the other three objects in the array) in 50 ms bins (Atkinson et al., 2018; Barr, 2008; Coco et al., 2016), and was weighted to control for eye movement-based dependencies (Barr, 2008). The empirical logits were submitted to a linear mixed effects model with fixed and random effects (including random slopes) fit in the same way as the accuracy models. In addition, we included a fixed effect of time to the polynomial order of two, allowing us to capture the potential non-linearity of fixations across time. Time 1 is the linear term of fixations across time, and Time 2 is the quadratic term, which can account for the curvature of fixations across time (a positive coefficient represents a decrease followed by an increase, and a negative coefficient represents an increase followed by a decrease). Additional model information is provided in the Supplementary Materials, Section 1b.

Results

Both models were analysed using R (R Core Team, 2018) and lme4 (Bates et al., 2015), and results include p-value estimates from the lmerTest package (Kuznetsova et al., 2017). The random effects structure for each model was maximally specified (Barr et al., 2013); if the maximal model did not converge, complexity was removed until convergence was achieved. See the notes below for the maximal model that converged for accuracy,[5] and the maximal model that converged for the eye movement data.[6]

Accuracy. Figure 2 shows the mean accuracy rate across speech rates. Less than one percent (i.e., 0.76%) of the responses did not fall within the four image regions and were eliminated. Participants showed relatively high accuracy overall, indicating good comprehension. General linear mixed effects models revealed no main effects or interactions (all $p > .34$).

Eye movements. As noted previously, the probe types were collapsed in the analysis. To ensure that there were no differences between them in terms of anticipatory eye movements, we added a fixed effect (probe type) to an initial model. We compared it to a model without (1) the main effect of probe type and (2) the interaction with probe type using a log likelihood ratio test. Neither the main effect of probe type, $\chi^2(1)=0.05$, $p=.81$, nor the interaction with probe type, $\chi^2(16)=8.94$, $p=.92$, improved fit. Therefore, all remaining analyses were collapsed across probe type. Figure 3 shows the empirical logit of fixations from 200 ms post verb onset to 1,200 ms post verb onset. Note the empirical logit was smoothed for visualisation purposes in the figures, but the raw empirical logit was used for data analysis. The model revealed a significant difference between speed 3.5 and 4.5 ($t=2.49$, $p<.05$), with looks to the target increasing from 3.5 to 4.5. There was a main effect of Time 1 ($t=3.90$, $p<.01$) with fixations increasing linearly across time. There was an interaction between speed 3.5 vs. 4.5 and age ($t=2.18$, $p<.05$), with older adults having a greater increase in fixations to the target across time relative to younger adults. There was also an interaction between 4.5 vs. 5.5 and age ($t=-2.07$, $p<.05$), with older adults showing a decrease in fixations to the target across time from speed 4.5 to 5.5 but younger adults showing an increase in fixations to the target. The model revealed an interaction between 5.5 vs. 6.0 and Time 1 ($t=3.62$, $p<.001$), with fixations to the target linearly decreasing across time from speed 5.5 to 6.0. None of the other results were significant. See Supplementary Materials, Section 1b for additional information, Table 2 for mean logit values (collapsed across time), and Figure S1 in the Supplementary Materials for fixation proportions.

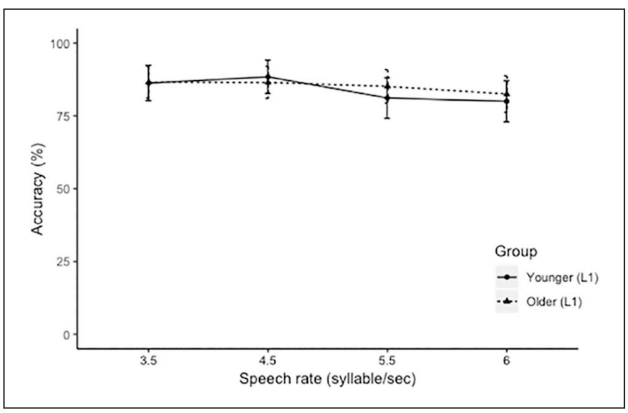

Figure 2. Mean accuracy for younger and older speakers. Error bars represent 95% confidence intervals.

Discussion

In the comparison between younger and older adults, it is clear that speech rate affected anticipatory eye movements for both groups but that there were marked differences between the groups. In all speech rates but the fastest (6.0 syllables per second), both younger and older adults show an empirical logit above 0 at the onset of the time window, suggesting that within 200 ms of the onset of the verb they are making anticipatory eye movements to the target. In the fastest condition, both groups made anticipatory eye movements to the target slightly after the onset of the time window. There was a significant increase in fixations to the target when moving from speed 3.5 to 4.5 for both groups, again suggesting that this increase facilitated anticipatory eye movements. In terms of age and speech rate, we found that older adults showed a greater increase in fixations to the target from speed 3.5 to 4.5 compared with younger adults. However, younger adults showed an

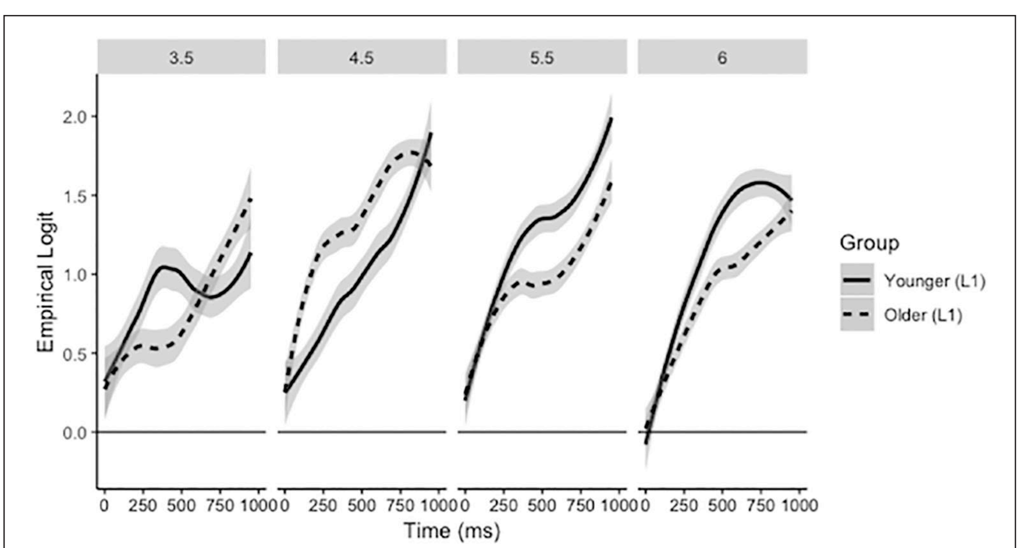

Figure 3. Empirical logit to the target object across time for younger and older (L1) speakers, the horizontal line indicates zero. Grey shading represents 95% confidence intervals.

increase in fixations up to speed 5.5, while older adults showed a decrease from 4.5 to 5.5. In addition, fixations to the target showed a linear decrease from 5.5 to 6.0.

To summarise, we hypothesised that as speech rate increased the number of anticipatory eye movements would decrease, and that this decrease would be greater in older adults. Instead, our results showed a more of an inverse-U pattern in which the number of anticipatory fixations increased from slower speech rates and then decreased with faster rates, suggesting an "optimal" speech rate in terms of anticipatory behaviour in the middle speech rate range. The implications and the mechanisms driving these results are discussed in more depth in the General Discussion.

Experiment 2

Experiment 2 compared young adults who were either native (L1) speakers of English or late L2 speakers of English (L1 German). Based on Experiment 1, we know that younger native speakers show the highest number of anticipatory fixations in the middle speech rate conditions. The key theoretical hypothesis for this experiment concerns the L2 speakers. Based on prior research, we hypothesised that L2 speakers would show no or very few anticipatory eye movements (i.e., they would more likely rely on heuristic processing), which is consistent with the Shallow Structure Hypothesis (Clahsen & Felser, 2006). Alternatively, if they do show anticipatory eye movements, they would do so only at the slowest speech rates, consistent with earlier reports of "delayed" anticipation (Dekydtspotter et al., 2006; see also Fernandez et al., 2018).

Method

Participants

English speakers (L1). The same sample of young adults from Experiment 1 were used in Experiment 2.

German speakers (L2). Twenty-three late L2 learners of English were recruited from the Technische Universität Kaiserslautern. All participants were born and currently lived in Germany, and no participant reported learning a second language before 6 years of age. Mean English proficiency score as measured by the Oxford Placement Test—Part A (OPT) was 40.60 out of 50. The L2 group had an average of 16.69 years of education (SD = 0.56), while the younger group had 15.00 years of education (SD = 0.89). This difference was significant using a Wilcoxon rank test W = 537.00, p < .05. In addition, there was a significant difference in age W = 553.50, p < .05. All had normal or corrected-to-normal vision and were paid 10 Euros. See Table 3 for participant information.

Materials. Same as Experiment 1.

Apparatus. Same as Experiment 1.

Procedure. The experimental session began with an information sheet and consent form. This was followed by the eye-tracking task, a language background questionnaire, an English grammar test to assess English proficiency (the Oxford Placement Test—Part A), and a working memory task (not presented here). The entire session lasted approximately 75 minutes.

Data analytic plan. Same as Experiment 1, except that the fixed effect of group was L1 vs. L2. See the notes below for the maximal model that converged for accuracy,[7] and the maximal model that converged for the eye movement data.[8] All code and data are available online at https://osf.io/9s7ye.

Results

Accuracy. Figure 4 shows the mean accuracy across the four speech rates. Less than 2% (1.98%) of the responses did not fall within the four image regions and were eliminated. General linear mixed effects models revealed a main effect of group (t = −4.91, p < .001), with L1 speakers having a greater accuracy compared with L2 speakers. There was also an interaction between condition 4.5 and 5.5 (t = 2.08, p < .05). Pairwise comparisons (using the *multcomp* package; Hothorn et al., 2008) revealed that L2 speakers were less accurate than L1 speakers in the 4.5 condition (b = 1.88, SE = 0.39, z = 4.82, p < .001), but did

Table 2. Mean empirical logit values by speech rate and age group.

Condition	Younger (L1)	Older (L1)
3.5	0.83 (3.21)	0.78 (3.11)
4.5	0.97 (3.05)	1.32 (2.97)
5.5	1.21 (3.17)	0.95 (3.09)
6.0	1.09 (3.12)	0.86 (3.07)

Table 3. L2 (German) participant information.

N	Mean age in years (*SD*)	Mean age of English acquisition (*SD*)	Gender	Mean OPT score (*SD*)	Mean years of education (*SD*)
23	25.03 (3.84)	10.39 (1.61)	10 female 13 male	40.60/50.00 (5.40)	16.69 (0.56)

SD: standard deviation; OPT: Oxford Placement Test.

not differ in accuracy in the 5.5 condition (b=0.75, SE=0.40, z=1.89, p=.55). There were no other main effects or interactions (all $ps > .10$); see Supplementary Materials, Section 2a for additional information.

Eye movements. Similar to Experiment 1, we ran log likelihood ratio tests with a model including a fixed effect of probe type, we compared this to a model without (1) a main effect of probe type and (2) the interaction of probe type. We found that the main effect of probe type improved the fit of the model, $\chi^2(1)=7.86$, $p < .01$, but the interaction with probe type did not, $\chi^2(16)=4.96$, $p=.99$. The main effect of probe type was such that the *point to* probes had fewer fixations to the target than the *wh*-word probes. However, given that the interaction with probe type did not influence model fit, we again collapsed across probe type. Figure 5 shows the empirical logit of fixations 200 ms following post verb onset to 1,000 ms post verb onset. The model revealed a main effect of group (t=2.92, $p < .01$) with L2 speakers showing significantly fewer fixations to the target compared with L1 speakers. There was a main effect of both Time 1 (t=3.12, $p < .01$) and Time 2 (t=−2.47, $p < .01$), suggesting that fixations showed a linear increase followed by a decrease across time. There was an interaction between speed 4.5 vs. 5.5 and Time 1, with fixations to the target showing a linear increase from speed 4.5 to 5.5 (t=2.03, $p < .05$). In addition, there was an interaction between 4.5 vs. 5.5 × group × Time 1 (t=−2.18, $p < .05$), and 5.5 vs. 6.0 × group × Time 1 (t=1.97, p=.05). From 4.5 to 5.5, there was a larger (linear) increase in fixations to the target by L1 speakers compared with L2 speakers. From 5.5 to 6.0, there was a linear decrease in fixations to the target across time for L1 speakers and no change in fixations by L2 speakers. See Supplementary Materials, Section 2b for additional information and Table 4 for mean empirical logit values (collapsed across time).

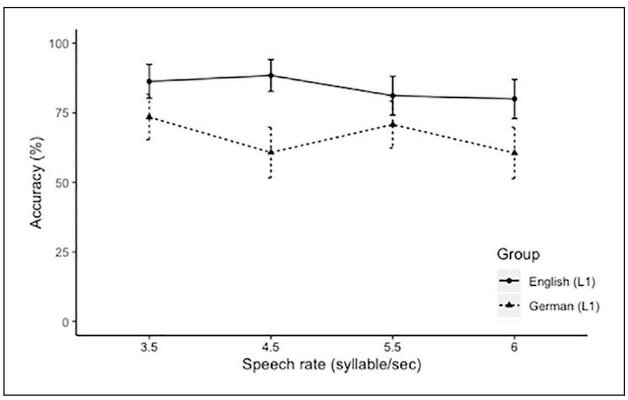

Figure 4. Mean accuracy for L1 and L2 speakers. Error bars represent 95% confidence intervals.

Discussion

Similar to Experiment 1, we observed that speech rate affected anticipatory eye movements for both groups, and that this impact was markedly different across groups. More specifically, L2 speakers are consistently later than L1 speakers in reliably making fixations to the target (L1 speakers begin at the latest by the onset of the time window). For L2 speakers, the point in time in which the empirical logit reliably occurs above zero varies (we define "reliably" as the first empirical logit value larger than zero for three or more consecutive time bins (150 ms)). For the 3.5 condition, the empirical logit crosses the zero line reliably starting in the third bin (100–150 ms), in the 4.5 condition in the eighth bin (350–400 ms), in the 5.5

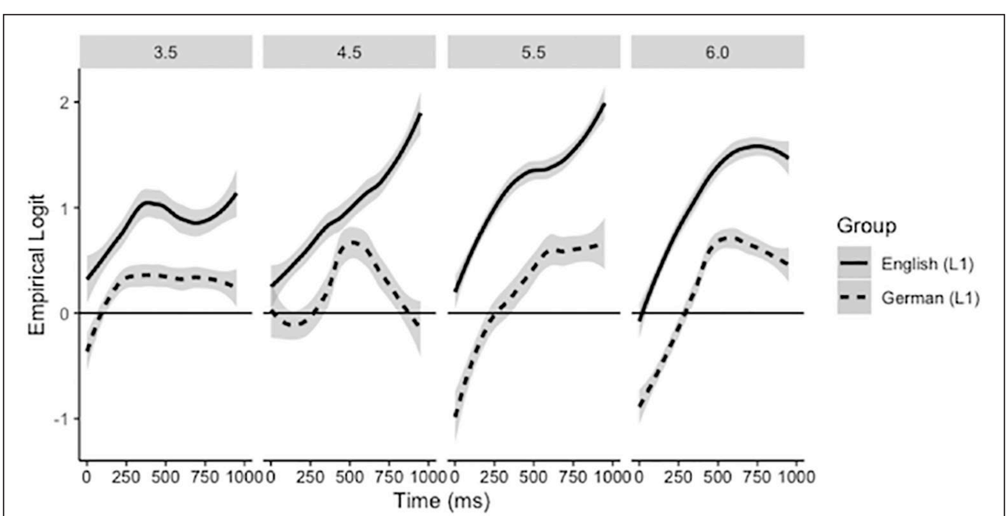

Figure 5. Empirical logit to the target object across time for L1 and L2 speakers, the horizontal line indicates zero. Grey shading represents 95% confidence intervals.

Table 4. Mean empirical logit values by speech rate and language group.

Condition	English (L1)	German (L2)
3.5	0.83 (3.21)	0.25 (3.29)
4.5	0.97 (3.05)	0.18 (3.01)
5.5	1.21 (3.17)	0.20 (3.15)
6.0	1.09 (3.12)	0.20 (3.19)

condition in the seventh bin (300–350 ms), and in the 6.0 condition in the eighth bin (350–400 ms).[9]

There were three main findings from Experiment 2. The first methodological finding was that fixations showed a similar inverse-U pattern to that found in Experiment 1, in which there were more anticipatory fixations with middle speech rates than with the slower and faster speech rates. This finding is not consistent with our hypothesis because we had expected a general decrease in anticipatory fixations in L2 speakers. The second finding, described above, was that L2 speakers showed delayed eye movements compared with L1 speakers. Moreover, the number of fixations on the target was always lower in L2 speakers. These findings are consistent with our theoretical hypotheses. The third main finding relates to the interactions, and here it is easiest to consider the groups separately. L1 speakers showed increased anticipatory eye movements from 4.5 to 5.5, but then a decrease from 5.5 to 6.0. In contrast, L2 speakers showed a smaller increase from 4.5 to 5.5 than L1 speakers, and then no difference from 5.5 to 6.0. Again, the final two findings are not consistent with our theoretical hypotheses, because we had expected a general decrease in anticipatory fixations in L2 speakers with increasing speech rate.

General discussion

This study had two main objectives: one methodological and one theoretical. The methodological objective focused on the role of speech rate on anticipatory eye movements. In the Introduction, we noted the lack of understanding of the role of speech rate as a major limitation (or flaw) in existing VWP research. In short, there is currently no standardised way of controlling or reporting the speech rate of auditory materials and, as a result, this critical information is omitted from virtually every paper. The obvious problem is that if speech rate affects participants' ability to generate anticipatory eye movements, then this may lead to conflicting findings, and clear over- or under-estimates of anticipatory behaviour. In this study, we have gone beyond labels of "normal" and "slow," providing quantitative data on our materials, which were manipulated for speech rate. In addition, our speech rates fell within the range of typical speech rates reported for conversational speech.

The theoretical objective focused on assessing the impact of speech rate on anticipatory eye movements in *older adults* and *L2 speakers*, two groups that may be differentially impacted by speech rate given prior reports on sentence processing. We investigated this by examining participants' ability to associate a filler with a gap site in filler-gap dependencies. Below, we first discuss the methodological findings from Experiment 1 and Experiment 2 together, and second, the theoretical findings. We then present the limitations and future directions, and we end with conclusions.

Methodological findings

The current study, like Huettig and Guerra (2019), showed differential patterns of anticipatory eye movements emerging across different speech rates. These findings reveal important insights about speech rate and the interaction between language, visual context, and shifts of attention (Ferreira & Tanenhaus, 2007; Henderson & Ferreira, 2004). At the most basic level, these finding show that there seems to be a "sweet spot" in speech rate, such that it leads to great synchronisation or coupling between language processing and shifts of attention to objects (which are indexed as anticipatory eye movements). By this explanation and in line with Ferreira and Chantavarin (2018), speech that is too fast does not permit incremental interpretations, such that the comprehension system cannot anticipate and then map that anticipation onto the available visual context. Likewise, speech that is too slow may permit very early anticipatory inferences, such that by the time the key word/phrase is heard, the eyes have already landed on and then moved from the target object. The "sweet spot" or peak in the inverse-U results when incremental comprehension and prediction lead to predictable time-locked (or coupled) shifts of attention to the target object (Altmann & Kamide, 2004).

The data from Experiment 1 suggest that older adults may have a "sweet spot" of 4.5 syllables per second (in line with Wingfield et al., 1985), while younger adults may have a "sweet spot" of 5.5 syllables per second (slightly faster than reported in Munro & Derwing, 2001). The data from Experiment 2 suggests that a speech rate of 4.5 syllables per second is right on the cusp of being too fast for L2 speakers (see Graham, 2006); they attend at this speed and shift their attention in line with the unfolding sentence (which is why we see the decrease in looks as time unfolds). In contrast, the 5.5 and 6.0 conditions are too fast; the L2 speakers wait for the entirety of verb to be uttered to commit to an interpretation, which leads to ultimately more correct interpretations. If this explanation is on the right track, then it is imperative that future VWP studies should, at the minimum, report speech rate and discuss the potential role of the speech rate used within the context of that study. This would make it easier to compare across studies, and to provide alternative explanations for competing results. The evidence presented in this study also suggests that the "ideal" speech rate may differ across groups/populations.

Theoretical findings

Our second objective was to test whether two groups thought to show fewer anticipatory eye movements (older L1 and younger L2 speakers) would be more impacted by speech rate compared with younger L1 speakers. Given shifting language processing abilities across the lifespan, we investigated native speakers of English across two age groups: younger and older adults. This study is the first to directly compare younger and older adults processing of wh-movement using the VWP. Both groups made anticipatory eye movements to the target in three speech rate conditions (3.5, 4.5, and 5.5). In the fastest condition (6.0), anticipatory eye movements to the target occurred around the onset of the time window (i.e., 200 ms following the onset of the verb). This suggests, in line with previous research (e.g., Dickey et al., 2007; Frazier & Flores D'Arcais, 1989; Sussman & Sedivy, 2003), that both groups actively searched for gaps and filled them at the earliest point possible.

Previous research suggests that older individuals develop strategies to compensate for age-related cognitive declines by (over) using context to build expectations (Rayner et al., 2006). The current study suggests that older adults behave similarly to younger adults, and even showed increased looks to the target from 3.5 to 4.5 syllables per second. This could be indicative of more proficient use of context to compensate for the increase in speed (in line with the risky reading strategy, Rayner et al., 2006). However, 5.5 syllables per second is too fast, and older are not able to compensate and showed fewer anticipatory eye movements relative to younger adults.

For L2 speakers, research has been mixed as to whether they are capable of forming filler-gap dependencies like L1 speakers, and whether they build expectations to the same extent as L1 speakers. Therefore, in Experiment 2, we investigated L1 speakers of English and L2 speakers of English (L1 German). We found that L2 speakers made significantly fewer fixations on the target than L1 speakers, and that L2 speakers had significantly lower comprehension accuracy than L1 speakers. Also, L2 speakers did not show reliable looks to the target until later than L1 speakers, and this timing varied across window sizes (see also Dekydtspotter et al., 2006). For L2 speakers, a faster speech rate led to decreased looks to the target in most cases (cf. comparison of 4.5 vs. 5.5).

In line with previous research, it seems that 3.5 syllables per second (and below) may be unnaturally slow and lead to unrepresentative eye movements for L1 speakers (Ferreira & Chantavarin, 2018; Huettig & Guerra, 2019). However, we argue that L2 speakers are anticipating upcoming objects in filler-gap dependencies, primarily at the slowest speed. This also provides evidence that when auditory presentation speed is on the lower end of "normal," L2 speakers can use top-down syntactic information and anticipate upcoming information and actively form filler-gap dependencies (see also Dekydtspotter et al., 2006; Fernandez et al., 2018; Hopp, 2006; Kaan, 2014; Pliatsikas & Marinis, 2013). However, when speech rate increases (particularly at 5.5 and 6.0), L2 speakers are unable to use syntactic information to posit a gap, but instead, rely more on the shallower route and wait to hear the verb before committing to an interpretation (Clahsen & Felser, 2018).

Limitations and future directions

We believe that the study had two main limitations. The first concerns the matching of groups. Our groups had different education and the two younger adult groups were different in age. Moreover, our data indicated that education was related to the dependent variables. The second limitation concerns the power of the study, which may be viewed on the low side. We encourage further research with more items per condition and larger sample sizes. As for future directions, we believe that it will be important to better understand which individual differences variables are associated with age-related declines in language processing (e.g., inhibition or working memory), and given that information, how compensatory strategies mitigate these declines. Second, VWP research would adopt standardised speech rates and reporting procedures. However, more work needs to be done with respect to speech rates to confirm optimal speech rates in anticipatory eye movement behaviour.

Conclusions

In summary and consistent with Huettig and Guerra (2019), we found evidence that speech rate impacted anticipatory eye movements in the processing of filler-gap dependencies. All three groups showed evidence of active gap filling, but did so at different speech rates. Native speakers benefitted most from speech rates between 4.5 and 5.5 syllables per second, with older adults showing more anticipatory eye movements at 4.5 syllables per second and younger adults at 5.5 syllables per second. L2 speakers benefitted most from speech rates at 3.5 syllables per second but may show increased benefit from a slightly faster speed (i.e., there was some evidence of increased fixations at 5.5). Our results highlight the methodological importance of controlling speech rate in VWP studies, as well as provide theoretical evidence that all groups are capable of forming filler-gap dependencies.

Acknowledgements

The authors thank Rebecca Jennings for recording the stimuli, Jasmine Segarra for assisting with collecting papers for the survey of VWP studies, and Sean Veld for assistance with the PRAAT script.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Leigh B Fernandez https://orcid.org/0000-0002-1635-2671

Data accessibility statement

The data from the present experiment are publicly available at the Open Science Framework website: https://osf.io/9s7ye

Supplementary material

The supplementary material is available at qjep.sagepub.com

Notes

1. *Acta Psychologica; Applied Psycholinguistics; Attention, Perception, & Psychophysics; Bilingualism: Language & Cognition; Brain & Language; Cognition; Journal of Memory and Language; Journal of Neurolinguistics; Language Acquisition; Language and Cognitive Processing; Language, Cognition, and Neuroscience; Linguistic Approaches to Bilingualism; Second Language Research; Studies in Second Language Acquisition; Quarterly Journal of Experimental Psychology*.
2. See Pickering and Barry (1991) for an argument against these types of movement operations, which will not be considered further in the current study.
3. Originally the auditory stimuli included a 6.5 syllables per second condition, but upon listening to the 6.5 conditions it was judged by the first author to be too fast, and therefore a 6.0 syllable per second condition was created instead.
4. Alternatively, one could analyse a set window (e.g., from the verb to the offset of the location); however, this would mean each time window would contribute different amounts of data points to the analyses, potentially leading to spurious results.
5. glmer(accuracy~speech_rate*group+(1+speech_rate|Participant)+(1+group*speech_rate|Item), family="binomial")
6. lmer(empirical_logit ~ speech_rate*group*poly(time,2)+(1+speech_rate+poly(time,2)|Participant)+(1+group+poly(time,2)|item
7. glmer(accuracy~speech_rate*group+(1+speech_rate|Participant)+(1+group+speech_rate|Item), family="binomial")
8. lmer(empirical_logit ~ speech_rate*group*poly(time,2)+(1+poly(time,2)|Participant)+(1 age+poly(time,2)|item))
9. Note that in the 4.5 condition the empirical logit falls below the zero line in the last two bins (900–1,000 ms)

References

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

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

Altmann, G. T. M., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action* (pp. 347–386). Psychology Press.

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

Atkinson, E., Wagers, M. W., Lidz, J., Phillips, C., & Omaki, A. (2018). Developing incrementality in filler-gap dependency processing. *Cognition*, *179*, 132–149. https://doi.org/10.1016/j.cognition.2018.05.022

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

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. (R Package Version 1.1-18-1). https://doi.org/10.18637/jss.v067.i01

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *352*(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

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

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*, 693–706. https://doi.org/10.1017/S0272263117000250

Coco, M. I., Keller, F., & Malcom, G. L. (2016). The role of visual context and visual memory. *Cognitive Science*, *40*, 1995–2024. https://doi.org/10.1111/cogs.12313

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.), *Proceedings of the 8th generative approaches to second language acquisition conference* (pp. 33–40). Cascadilla Proceedings Project.

Dickey, M. W., Choy, J. J., & Tompson, C. K. (2007). Real-time comprehension of wh-movement in aphasia: Evidence from

eyetracking while listening. *Brain and Language*, *100*, 1–22. https://doi.org/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 eye tracking. *Journal of Neurolinguistics*, *22*, 563–583. https://doi.org/10.1016/j.jneuroling.2009.06.004
Dussias, P. E., Valdés, J. R. K., Guzzardo, R. E. T., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, *35*, 353–387. https://doi.org/10.1017/S0272263112000915
Fernandez, L. B., Höhle, B., Brock, J., & Nickels, L. (2018). Investigating auditory processing of syntactic gaps with L2 speakers using pupillometry. *Second Language Research*, *34*, 201–227. https://doi.org/10.1177/0267658317722386
Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, *27*, 443–448. https://doi.org/10.1177/0963721418794491
Ferreira, F., & Tanenhaus, M. K. (2007). Introduction to the special issue on language-vision interactions. *Journal of Memory and Language*, *57*, 455–459. https://doi.org/10.1016/j.jml.2007.08.002
Fónagy, I., & Magdics, K. (1960). Speed of utterance in phrases of different lengths. *Language and Speech*, *3*, 179–192. https://doi.org/10.1177/002383096000300401
Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory*, *5*, 519–559. https://doi.org/10.1007/BF00138988
Frazier, L., & Flores D'Arcais, G. B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*(3), 331–344. https://doi.org/0.1016/0749-596X(89)90037-5
Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax, 7*, 55–78. https://doi.org/10.1111/j.1368-0005.2004.00065.x
Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, *34*, 165–182. https://doi.org/10.1016/j.system.2005.11.001
Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.
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. https://doi.org/10.1177/0267658312437990
Grüter, T., Rohde, H., & Schafer, A. J. (2014). The role of discourse-level expectations in non-native speakers' referential choices. In W. Orman & M. J. Valleau (Eds.), *Proceedings of the 38th Boston University conference on language development* (pp. 179–191). Cascadilla Press.
Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, *7*(2), 199–229. https://doi.org/10.1075/lab.15011.gru
Henderson, J., & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world*. Psychology Press.
Hertrich, I., Dietrich, S., & Ackermann, H. (2013). How can audiovisual pathways enhance the temporal resolution of time-compressed speech in blind subjects? *Frontiers in Psychology*, *4*, e530. https://doi.org/10.3389/fpsyg.2013.00530
Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22*, 369–397. https://doi.org/10.1191/0267658306sr272oa
Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*, 33–56. https://doi.org/10.1177/0267658312461803
Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*, 346–363.
Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014
Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research*, *1706*, 196–208. https://doi.org/10.1016/j.brainres.2018.11.013
Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, *31*(1), 19–31. https://doi.org/10.1080/23273798.2015.1072223
Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003
Kaan, E. (2014). Predictive sentence processing in L2 and L1 what is different? *Linguistic Approaches to Bilingualism*, *4*, 267–282. https://doi.org/10.1075/lab.4.2.05kaa
Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, *2*(4), 647–670. https://doi.org/10.1111/j.1749-818X.2008.00072.x
Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*, 95–127. https://doi.org/10.1016/j.cognition.2004.03.002
Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26. (R Package Version 3.0-1). https://doi.org/10.18637/jss.v082.i13
Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*, 133–156. https://doi.org/10.1080/03640210709336987
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*, 53–78. https://doi.org/10.1017/S027226310505003
Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate. *Studies in Second Language Acquisition*, *23*, 451–468.
Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-

up and top-down information processing. *International Journal of Audiology*, *47*, S72–S82. https://doi.org/10.1080/14992020802307404

Pickering, M. J., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, *6*, 229–259. https://doi.org/10.1080/01690969108406944

Pliatsikas, C., & Marinis, T. (2013). Processing empty categories in second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition*, *16*, 167–182. https://doi.org/10.1017/S136672891200017X

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*, 448–465. https://doi.org/10.1037/0882-7974.21.3.448

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rosselli, M., Ardila, A., Matute, E., & Vélez-Uribe, I. (2014). Language development across the life span: A neuropsychological/neuroimaging perspective. *Neuroscience Journal*, *2014*, 585237. https://doi.org/10.1155/2014/585237

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychology Society*, *16*, 754–760. https://doi.org/10.1017/S1355617710000706

Schik, G., Mohr, S., & Hofferberth, B. (2000). Effects of aging on saccadic eye movements to visual and auditory targets. *International Tinnitus Journal*, *6*, 154–159.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual interpretation. *Cognition*, *71*, 109–147. https://doi.org/10.1016/S0010-0277(99)00025-6

Shafto, M. A., & Tyler, L. K. (2014). Language in the aging brain: The network dynamics of cognitive decline and preservation. *Science*, *346*, 583–586. https://doi.org/10.1126/science.1254404

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language & Linguistics Compass*, *9*, 311–327. https://doi.org/10.1111/lnc3.12151

Stine-Morrow, E. A. L., Miller, L. M. S., & Hertzog, C. (2006). Aging and self-regulated processing. *Psychological Bulletin*, *132*, 582–606. https://doi.org/10.1037/0033-2909.132.4.582

Sussman, R. S., & Sedivy, J. C. (2003). The time course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes*, *18*, 143–161. https://doi.org/10.1080/01690960143000498

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634. https://doi.org/10.1126/science.7777863

Warrington, K. L., White, S. J., & Paterson, K. B. (2018). Ageing and the misperception of words: Evidence from eye movements during reading. *Quarterly Journal of Experimental Psychology*, *71*, 75–84. https://doi.org/10.1080/17470218.2016.1251471

Wilshire, C. (1999). The "tongue twister" paradigm as a technique for studying phonological encoding. *Language and Speech*, *42*, 57–82. https://doi.org/10.1177/00238309990420010301

Wingfield, A., Poon, L. W., Lombardi, L., & Lowe, D. (1985). Speed of processing in normal aging: Effects of speech rate, linguistic structure, and processing time. *Journal of Gerontology*, *40*, 579–585. https://doi.org/10.1093/geronj/40.5.579

Zurif, E., Swinney, D., Prather, P., Wingfield, A., & Brownwell, H. (1995). The allocation of memory resources during sentence comprehension: Evidence from the elderly. *Journal of Psycholinguistic Research*, *24*, 165–182. https://doi.org/10.1007/BF02145354