Vocabulary knowledge predicts individual differences in the integration of visual and linguistic constraints

Anuenue Kukona, Olivia Gaziano, Marie-Josee Bisson and Adrian Jordan

Division of Psychology, De Montfort University

Author Note

Anuenue Kukona https://orcid.org/0000-0003-4377-3057

Correspondence concerning this article should be addressed to Anuenue Kukona, Division of Psychology, De Montfort University, The Gateway, Leicester, LE1 9BH, UK. Telephone: +44 (0)116 250 6184 Email: anuenue.baker-kukona@dmu.ac.uk

The Version of Record of this manuscript has been published and is available in Language, Cognition and Neuroscience http://www.tandfonline.com/10.1080/23273798.2021.2020854.

Abstract

Two experiments investigated individual differences in the integration of visual and linguistic constraints during syntactic ambiguity resolution. Skilled adult comprehenders heard sentences like “Put the kiwi on the rectangle on the circle”, in which “on the rectangle…” could temporarily reflect either the destination of “put” or a modifier of “kiwi”, while viewing visual arrays with either 1 kiwi (e.g., on a rectangle) or 2 kiwis (e.g., on a rectangle vs. triangle). While the noun “kiwi” provided sufficient information to distinguish the object of interest in the 1 referent context, modification was necessitated by the 2 referent context. Garden path eye (Experiment 1) and mouse (Experiment 2) movements to the incorrect (e.g., rectangle) destination were reduced in 2 vs 1 referent contexts, conceptually replicating prior findings, and these effects were weaker for participants with less vs. more vocabulary knowledge. Implications for models of sentence processing are discussed.

Keywords: individual differences; language experience; syntactic ambiguity resolution; visual world paradigm; vocabulary knowledge

Introduction

Sentence comprehension draws on both linguistic and non-linguistic sources of information (e.g., Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). However, individual differences in sentence processing, particularly among adults, are poorly understood. The aim of the current research was to explore systematic variation in the integration of visual and linguistic constraints among skilled adult comprehenders.

Tanenhaus et al. (1995) investigated the influence of (i.e., non-linguistic) visual constraints on sentence comprehension. Their participants heard temporarily ambiguous sentences like “Put the apple on the towel in the box”, in which “on the towel…” could temporarily reflect either the destination of “put” or a modifier of “apple”, while viewing visual arrays with either 1 apple (e.g., on a towel) or 2 apples (e.g., on a towel vs. napkin), alongside an empty towel and other distractor objects. In the 1 referent context, the noun “apple” provided sufficient information to distinguish the object of interest, and thus modification was superfluous. In the 2 referent context, the noun “apple” provided insufficient information to do so, necessitating modification (e.g., such that “on the towel” provided necessary information about the apple of interest). Tanenhaus et al. recorded participants’ eye movements and observed a clear garden path effect in the 1 referent context: participants made eye movements to the empty towel on the majority of trials, suggesting that they incorrectly parsed “on the towel…” as the destination of “put”. In contrast, participants made significantly fewer eye movements to the empty towel in 2 referent contexts, suggesting that they correctly parsed “on the towel…” as a modifier, as necessitated by this context.

Tanenhaus et al.’s (1995) findings provide compelling evidence against modularity (e.g., Fodor, 1983). For example, garden path theory (e.g., Frazier & Rayner, 1982) hypothesises that parsing is initially dependent on syntactic information but independent of other constraints, reflecting an informationally encapsulated module. For a syntactically ambiguous sentence like “The second wife will claim the inheritance belongs to her”, in which “the inheritance…” could temporarily reflect either a direct object or sentential complement, garden path theory assumes that this phrase is initially parsed as the former based on syntactic constraints (e.g., minimal attachment). Consistent with this hypothesis, Frazier and Rayner observed slowed read times at disambiguation (e.g., “belongs…”), which necessitated reanalysis as the latter. In contrast, Tanenhaus et al.’s findings support interactive approaches, which assume that non-syntactic information, and even the visual context, can immediately constrain sentence comprehension (e.g., MacDonald, Pearlmutter & Seidenberg, 1994; McRae, Spivey-Knowlton & Tanenhaus, 1998; Trueswell & Tanenhaus, 1994).

On the one hand, visual world research emphasises comprehenders’ striking sensitivity to subtle visual minutia. For example, Chambers, Tanenhaus and Magnuson (2004) investigated object affordances. Their participants heard syntactically ambiguous sentences like “Pour the egg in the bowl over the flour”, while viewing visual arrays with either 2 eggs in liquid form (e.g., in a bowl vs. cup) or 1 egg in liquid form (e.g., in a bowl) and 1 egg in solid form (e.g., in a cup), alongside an empty bowl and other distractor objects. In both contexts, the noun “egg” provided insufficient information to distinguish the object of interest. However, the affordances of the egg in solid form were not verb relevant (e.g., pourable, in contrast to the eggs in liquid form), and thus modification was superfluous in this context. Similar to a 1 referent context, Chambers et al. found that participants made more eye movements to the incorrect destination (e.g., the empty bowl) with the egg in solid form, suggesting that object affordances immediately influenced parsing. In related research, Coco and Keller (2015) also found that the visual salience of objects influenced parsing. After hearing “Put…” in related syntactically ambiguous sentences, their participants made more eye movements to visually salient objects vs. controls (e.g., a vibrant vs. dull bowl), suggesting that they were anticipating potential verb arguments based on their saliency. Taken together, these findings suggest that sentence comprehension is immediately and robustly constrained by the visual context.

On the other hand, limitations on the integration of visual and linguistic constraints have also been observed. For example, Ferreira, Foucart and Engelhardt (2013) manipulated the preview and complexity of their visual contexts. Contrasting with Tanenhaus et al. (1995), when they eliminated the previews of their visual arrays, and when they tripled the number of objects in their visual arrays, they found that participants' eye movements to the incorrect destination did not differ between syntactically ambiguous 1 and 2 referent contexts. To the contrary, these findings suggest that sentence comprehension may not always be constrained by the visual context.

Likewise, developmental differences have also been observed. For example, Trueswell, Sekerina, Hill and Logrip (1999) found that five year olds' eye movements to the incorrect destination did not differ between 1 and 2 referent contexts, contrasting with adults (e.g., Tanenhaus et al., 1995). In addition, children made performance errors (i.e., while enacting the corresponding instructions) on the majority of syntactically ambiguous trials. Woodard, Pozzan and Trueswell (2016) replicated this pattern, and also found that children's performance errors were (inversely) predicted by their executive function (but not working memory or grammatical) skills. However, in contrast to their performance errors, children's eye movements to the incorrect destination were not predicted by these measures. In related computer mouse tracking research, Anderson, Farmer, Goldstein, Schwade and Spivey (2011) also found that children's performance errors were (inversely) predicted by their vocabulary knowledge. Taken together, these findings suggest that the visual context represents a still developing constraint on sentence comprehension at five years of age (e.g., see also Qi, Love, Fisher & Brown-Schmidt, 2020; Snedeker & Trueswell, 2004).

Visual world research has also begun to explore individual differences among adults. A handful of studies have compared the integration of visual and linguistic constraints in native and non-native speakers. For example, Pozzan and Trueswell (2016) found that the eye movements patterns of native English speakers did not differ from Italian L2 speakers of English: both groups made more eye movements to the incorrect destination in 1 vs. 2 referent contexts. However, in contrast to these eye movement patterns, non-native speakers' performance errors were comparable to five year olds. In related research, native English speakers also showed similar eye movement patterns to Spanish (Contemori, Pozzan, Galinsky & Dussias, 2018) and Japanese (Nakamura, Arai, Hirose & Flynn, 2020) L2 speakers of English. On balance, these findings suggest that non-native speakers are as sensitive to the visual context as native speakers are. However, in contrast to native speakers, non-native speakers struggle considerably with reanalysis (i.e., as reflected in their performance errors). Nakamura et al. suggest that non-native speakers may be especially sensitive to visual cues because they are salient and language independent. Finally, Rabagliati, Delaney-Busch, Snedeker and Kuperberg (2019) found that people with schizophrenia were not influenced by the visual context, although their performance errors were comparable to controls. Rabagliati et al. suggest that top-down integration processes may be specifically impaired in people with schizophrenia.

In summary, visual world research provides important insight into the influence of the visual context on sentence comprehension (e.g., among many other phenomena; for review, see Huettig, Rommers & Meyer, 2011; Knoeferle & Guerra, 2016; Salverda & Tanenhaus, 2017). However, the literature paints a mixed picture: while some research suggests that comprehenders are strikingly sensitive to the visual context (e.g., Chambers et al., 2004; Coco & Keller, 2015), and that this pattern is stable across different groups (e.g., see the eye movement patterns in Contemori et al., 2018; Nakamura et al., 2020; Pozzan & Trueswell, 2016), other research emphasises comprehenders' varying (in)sensitivity (e.g., Anderson et al., 2011; Ferreira et al., 2013; Rabagliati et al., 2019; Trueswell et al., 1999; Woodard et al., 2016). The aim of the current research was to investigate individual differences in the integration of visual and linguistic constraints among skilled adult comprehenders. While Woodard et al. (2016) and Anderson et al. (2011) have addressed cognitive predictors of children's comprehension, whether these patterns generalise to adults remains unclear. Moreover, this research emphasises individual differences in children's performance errors, which reflect late garden path reanalysis processes, rather than their eye movement patterns, which reflect immediate constraint integration processes. Finally, whether the qualitatively distinct eye movement patterns observed by Trueswell et al. (1999), Rabagliati et al. (2019) and others with qualitatively distinct groups (e.g., children vs. adults; people with schizophrenia vs. controls) exist along a continuum also remains unclear.

Motivated by a growing literature, the current research focused on individual differences related to vocabulary knowledge. First, Anderson et al.'s (2011) findings in children motivate a specific link between vocabulary knowledge and constraint integration processes. On their account, vocabulary knowledge provides an index of linguistic experience, such that children with greater linguistic experience made fewer errors because they were more adult like. However, to the extent that adults also vary in their vocabulary knowledge, and by proxy their linguistic experience, this account may also predict a related pattern in adults. Second, individual differences in vocabulary knowledge have also been linked to sentence processing in adults. For example, vocabulary knowledge is a compellingly persistent predictor of sentence level phenomena like prediction (e.g., Borovsky, Elman & Fernald, 2012; Hintz, Meyer & Huettig, 2017; Kukona et al., 2016; Rommers, Meyer & Huettig, 2015) and has also been linked to the processing of syntactic ambiguities (e.g., Engelhardt, Nigg & Ferreira, 2017). Likewise, while the sentence processing literature has classically emphasised working memory (e.g., Just & Carpenter, 1992), Van Dyke, Johns and Kukona (2014) found that among a comprehensive battery of individual differences measures, only vocabulary knowledge uniquely predicted readers' comprehension of syntactically complex sentences in a dual task paradigm. Perfetti (2007) articulates a Lexical Quality Hypothesis that emphasises the role of high quality lexical representations, which are characterised by detailed representations of form and rich representations of meaning, in (e.g., reading) comprehension. On this account, lesser quality lexical representations may impair sentence processing (e.g., when lexical representations are of poor quality, the sentential representations that are built of them may also be of poor quality). Thus, to the extent that adults also vary in their vocabulary knowledge, and by proxy the quality of their lexical representations, this account may also predict impairments in constraint integration processes in adults.

In two experiments, which were closely based on Tanenhaus et al. (1995), undergraduate participants heard syntactically ambiguous sentences like "Put the kiwi on the rectangle on the circle", while viewing visual arrays like Figure 1, with either 1 kiwi (e.g., on a rectangle) or 2 kiwis (e.g., on a rectangle vs. triangle), alongside other distractor objects. While modification of "kiwi" was superfluous in the 1 referent context, it was necessitated by the 2 referent context. In Experiment 1, like Tanenhaus et al. (1995), participants' eye movements were recorded. In Experiment 2, to conceptually replicate Experiment 1 and explore the sensitivity of mouse tracking, participants' mouse movements were recorded. If vocabulary knowledge is a determinant of these constraint integration processes, we predicted that comprehenders with more vocabulary knowledge would show the clearest divergence between 1 vs. 2 referent contexts (e.g., as reflected in eye and mouse movements to the incorrect [e.g., rectangle] destination), while comprehenders with less vocabulary knowledge would show a significantly weaker pattern.

Experiment 1

Method

In order to investigate individual differences in the integration of visual and linguistic constraints, participants completed a standardised test of vocabulary knowledge (Dunn & Dunn, 2007) and a visual world experiment in which they heard ambiguous sentences like "Put the kiwi on the rectangle on the circle" (i.e., vs. unambiguous controls) while viewing visual arrays with either 1 or 2 kiwis (e.g., see Tanenhaus et al., 1995). In Experiment 1, participants' eye movements were recorded.

Participants

Forty seven undergraduates (age M = 19.40, SD = 1.15) from De Montfort University participated for course credit. All participants were native English speakers with normal or corrected-to-normal vision. A power analysis using a correlation approach (e.g., relating sentence processing and vocabulary) was run based on the availability of both effect size estimates and well-established methods. An analysis using the *pwr* package in *R* revealed that the sample enabled detection of an effect size (r = 0.40, power = .80, α = .05) comparable to Qi et al. (2020), who observed a correlation of r = 0.42 between children's executive function and performance errors. Alongside correlational analyses, mixed effects analyses are also reported, which address the experimental manipulations directly. The sample was also consistent with prior individual differences research (e.g., N = 43, Anderson et al., 2011; N = 40, Novick, Thompson-Schill & Trueswell, 2008; N = 40, Woodard et al., 2016). The study was conducted with formal ethics approval and participants gave written consent.

Design

Both referent context (1 and 2) and instruction type (ambiguous and unambiguous) were manipulated within participants.

Materials

Twenty four experimental visual arrays were created using images from the MultiPic stimulus set (Duñabeitia, Crepaldi, Meyer, New, Pliatsikas, Smolka & Brysbaert, 2018). Each visual array (see Figure 1) included a pink circle, rectangle and triangle in the lower visual array, which reflected the source locations, and a gray circle, rectangle and triangle in the upper visual array, which reflected the destination locations. The source and destination locations were constant (e.g., within and between participants). Visual displays were 1,920 x 1,080 pixels and images were 300 x 300 pixels. Visual displays were divided into 2 x 3 visual arrays with six equally sized (i.e., 640 x 540 pixels) regions (i.e., which also corresponded to the eye movement interest areas), with a source or destination location centred within each region. At the beginning of each trial, a visual object was depicted at each source location, one of which was the target. In the 1 referent context, all three visual objects were different (e.g., kiwi, shirt and television). In the 2 referent context, the target and one other visual object were identical (e.g., kiwi, kiwi and television). Mirroring Tanenhaus et al., each visual array was also presented with a corresponding syntactically ambiguous (e.g., “Put the kiwi on the rectangle on the circle”) or unambiguous (e.g., “Put the kiwi that’s on the rectangle on the circle”) sentence. The source and destination locations of the targets were counterbalanced across items. Sentences were recorded by a female native speaker of British English. The mean duration of sentences was 3,221 ms (*SD* = 240) and their peak amplitudes were normalised. The full list of experimental items is reported in Table A1 in the

Appendix.

Four counterbalanced lists were created by dividing the 24 visual arrays into four groups and rotating them through the four (referent context x instruction type) conditions. Each list included six visual arrays per condition and each visual array appeared once on each list and once in each condition across lists. Twenty four filler trials were also created that included similar visual arrays but simpler sentences (e.g., “Put the mirror on the rectangle”), which only referred to destination and not source locations. One half of filler visual arrays also included two identical visual objects, but the target was never among these. Participants were randomly assigned to lists.

Finally, (i.e., receptive) vocabulary knowledge was measured using the final 60 items of the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007). Participants completed all 60 items and their performance was measured as their (i.e., raw) number of accurate responses out of 60. PPVT orders items from least to most difficult and has been normed with both children and adults. It was anticipated that this sample of undergraduates would show maximal variability among the most difficult (i.e., final) items. As anticipated, accuracies among these items spanned from approximately chance (i.e., 25%) to just below ceiling (Experiment 1: min = 30%; max = 73%; Experiment 2: min = 22%; max = 78%), reflecting considerable individual differences, even in this skilled sample. However, caution may be required when comparing these scores to other studies.

Procedure

Participants were tasked with moving one visual object on each trial from a pink source to gray destination location. Participants used the mouse to click on the visual object and drag it from its source to its destination location. Each trial was accompanied by a sentence that identified the visual object and destination location of interest (i.e., experimental sentences also identified the source location), which was presented via speakers. Participants began each trial by clicking on a central fixation cross, and then they previewed the visual arrays, such that the onset of the visual array preceded the onset of the sentence by 4,000 ms. Trials ended after mouse click release. Participants’ eye movements were tracked throughout using a Tobii Nano Pro sampling at 60 Hz. The experiment included 24 experimental trials, equally distributed across the four (i.e., referent context x instruction type) conditions. The experiment also included two practice and 24 filler trials. The order of trials was pseudorandomised. Finally, participants completed the vocabulary knowledge test after the experiment.

Results

Accuracy

Mean accuracy was 95.39% (*SD* = 11.88) in the 1 referent and ambiguous instruction condition, 99.64% (*SD* = 2.43) in the 1 referent and unambiguous instruction condition, 95.74% (*SD* = 7.34) in the 2 referent and ambiguous instruction condition and 97.87% (*SD* = 5.62) in the 2 referent and unambiguous instruction condition. Mean accuracies were submitted to a by-participants (i.e., reflecting the current focus on a participant-level predictor) mixed effects model with fixed effects of referent context (1 = -0.5; 2 = 0.5), instruction type (ambiguous = -0.5; unambiguous = 0.5) and vocabulary knowledge (z-score; *M* = 33.23, *SD* = 5.96), alongside their interactions, and random intercepts by participants (*lme4*: “Accuracy ~ Ref * Instr * Vocab + (1|Participant)”; Bates, Mächler, Bolker & Walker, 2014). For this by-participants analysis, the outcome was generated by averaging over items/trials, which yielded a single data point (i.e., mean) per participant per condition; thus, random participant slopes (and random item effects) were not included in the model. Results are reported in Table 1. Accuracy was significantly lower for ambiguous vs. unambiguous instructions, and marginally lower for participants with lower vs. higher vocabulary knowledge.

(Table 1 about here)

Eye movements

Inaccurate trials were excluded from the eye movement analyses. Average proportions of fixations to target (e.g., the kiwi on the pink rectangle with “Put the kiwi on the rectangle on the circle”) and competitor/distractor (e.g., the kiwi/shirt on the pink triangle) referents and correct (e.g., gray circle) and incorrect (e.g., gray rectangle) destinations are plotted in Figures 2 and 3. The plots span sentence onset to 600 ms following sentence offset, with eye movements (re)synchronised at sentence onset and the onset and offset of each noun (e.g., “kiwi”, “rectangle” and “circle”). As expected, fixations to the target and distractor referents in the 1 referent context diverged shortly after the first noun (e.g., “kiwi”), such that modification of the target was superfluous, while fixations to the target and competitor referents in the 2 referent context diverged only after the second noun (e.g., “rectangle”), such that modification was necessitated by the 2 referent visual context.

In order to assess garden path effects across conditions (i.e., misanalysis of “on the rectangle” as a destination), eye movements to the incorrect destination (e.g., see Tanenhaus et al., 1995; gray rectangle) between the offset of the second noun and sentence offset (e.g., “…on the circle”) were analysed. Mean proportions of fixations are plotted by condition in Figure 4A. Mean proportions of fixations were arcsine square root transformed and submitted to a by-participants mixed effects model with fixed effects of referent context, instruction type and vocabulary knowledge and random intercepts by participants. Results are reported in Table 2 and depicted in Figure 4B.

(Table 2 about here)

Eye movements to the incorrect destination were significantly lower in the 2 vs. 1 referent context, and they were significantly lower for unambiguous vs. ambiguous instructions. In addition, the significant interaction of referent context and instruction type revealed that the effect of referent context was more pronounced for ambiguous vs. unambiguous instructions. Finally, vocabulary knowledge interacted significantly with referent context, such that the effect of referent context was more pronounced for participants with more vs. less vocabulary knowledge. In other words, the divergence between 1 referent (i.e., upward orientated symbols; see Figure 4B) vs. 2 referent (i.e., downward orientated symbols) contexts was significantly weaker for participants with less (i.e., on the left) vs. more (i.e., on the right) vocabulary knowledge. Similarly, a correlational analysis of the difference in eye movements between 1 vs. 2 referent contexts (i.e., collapsing over instruction type) revealed a significant correlation with vocabulary knowledge, $r(45) = 0.38$, $p < 0.01$. In contrast, a correlational analysis of the difference in eye movements between ambiguous vs. unambiguous instructions (i.e., collapsing over referent context) revealed a non-significant correlation with vocabulary knowledge, $r(45) = 0.07$, $p = 0.65$.

Discussion

Conceptually replicating prior findings (e.g., Tanenhaus et al., 1995), participants in Experiment 1 made less garden path eye movements to the incorrect destination in the 2 referent visual context. In addition, the divergence between 2 vs. 1 referent visual contexts was significantly weaker for participants with less vs. more vocabulary knowledge, linking constraint integration processes to vocabulary knowledge.

Alongside eye movements, mouse movements have also been used to investigate constraint integration processes (e.g., Anderson et al., 2011; Farmer, Anderson & Spivey, 2007; Farmer, Cargill, Hindy, Dale & Spivey, 2007). For example, using stimuli like Tanenhaus et al. (1995), Farmer, Cargill et al. (2007) found that the trajectories of participants' mouse movements were more attracted to incorrect destinations (e.g., an empty towel when hearing "Put the apple on the towel in the box") in 1 vs. 2 referent contexts, suggesting that they incorrectly parsed the ambiguous phrase as a destination during the former (i.e., garden pathed).

The aim of Experiment 2 was to both conceptually replicate the influence of vocabulary knowledge on constraint integration processes and explore the sensitivity of mouse movements to individual differences in these processes. Experiment 2 paralleled Experiment 1 except that participants' mouse movements were recorded. Again, if vocabulary knowledge is a determinant of constraint integration processes, we predicted that comprehenders with more vocabulary knowledge would show the clearest divergence between 1 vs. 2 referent contexts.

Experiment 2

Method

In order to investigate individual differences in the integration of visual and linguistic constraints, participants completed an online study that paralleled Experiment 1. In Experiment 2, participants' mouse (i.e., rather than eye) movements were recorded.

Participants

Forty seven undergraduates (age *M* = 19.79, *SD* = 2.66) from De Montfort University participated for course credit. All participants were native English speakers with normal or corrected-to-normal vision who did not take part in Experiment 1. Based again on a correlational approach, the sample enabled detection of an effect size comparable to Qi et al. (2020) and Experiment 1, in which a correlation of $r = 0.38$ was observed between participants' vocabulary knowledge and eye movements in 1 vs. 2 referent contexts. Five participants were excluded whose data included timing errors (i.e., non-monotonic time samples) on a majority of their trials.

Design

As in Experiment 1, both referent context (1 and 2) and instruction type (ambiguous and unambiguous) were manipulated within participants.

Materials

The sentence materials (e.g., including practice and filler sentences) were identical to Experiment 1. However, the visual arrays from Experiment 1 were simplified to constrain participants' patterns of mouse movements and enable these to be aggregated across trials in the analyses. For example, mouse movements in Experiment 1 (see Figure 1) to the gray triangle destination from the pink circle source spanned a farther distance than from the pink rectangle source, a problem for aggregating across such movements. Visual arrays included two rather than three visual objects (i.e., and source and destination locations; see Figure 5). For the 1 referent context, the distractor visual object (e.g., shirt) was not included in the visual array, while the target visual object (e.g., kiwi) and the incorrect destination shape (e.g., square) were horizontally positioned on one side of the visual array and the other visual object (e.g., television) and the correct destination shape (e.g., circle) were positioned on the opposite side. For the 2 referent context, the 1 referent context was modified to include two identical visual objects (i.e., both reflecting the target visual object). As an online study, participants' visual displays (e.g., screen resolutions, etc.) varied. Visual arrays used normalised units that ranged from -1 to 1. Images were 0.3 x 0.6 normalised units (i.e., images appeared perfectly square for a 2:1 screen ratio and compressed/stretched for other ratios), and they were positioned in the extreme corners of the visual array. Participants were always required to move target visual objects from correct source locations on one side of the visual array to correct destination locations on the opposite side of the visual array, simplifying their mouse trajectories and allowing the horizontal axis to be inverted for target visual objects on the left vs. right side of the visual array. The location of target visual objects and (source/destination) shapes (i.e., on the left vs. right side of the visual array) was counterbalanced across items. The full list of experimental items is reported in Table A1 in the Appendix.

Participants were randomly assigned to the same counterbalanced lists as Experiment 1. Vocabulary knowledge was again measured using the final 60 items of the Peabody Picture Vocabulary Test.

Procedure

The procedure was similar to Experiment 1 but conducted online using Pavlovia (https://www.pavlovia.org) and Qualtrics (https://www.qualtrics.com). However, the preview was reduced to 3,000 ms for the simplified visual arrays in Experiment 2, although this is equivalent to the preview in Ferreira et al. (2013). Participants' mouse movements were tracked throughout using Pavlovia. For the vocabulary knowledge test, participants played recordings of the words (i.e., alongside the pictures) using Qualtrics.

Results

Participants' vocabulary knowledge (M = 35.07, SD = 8.21) did not differ significantly from Experiment 1, $t(87) = -1.23$, $p = 0.23$.

Accuracy

Trials with timing errors (i.e., non-monotonic time samples) were excluded from all analyses (<1% of trials). Mean accuracy was 91.98% (SD = 13.95) in the 1 referent and ambiguous instruction condition, 92.30% (SD = 13.72) in the 1 referent and unambiguous instruction condition, 90.87% (*SD* = 13.38) in the 2 referent and ambiguous instruction condition and 92.46% (*SD* = 13.87) in the 2 referent and unambiguous instruction condition. Mean accuracies were submitted to a by-participants mixed effects model with fixed effects of referent context, instruction type and vocabulary knowledge and random intercepts by participants (i.e., identical to Experiment 1). Results are reported in Table 3. Accuracy was significantly lower for participants with lower vs. higher vocabulary knowledge.

(Table 3 about here)

Mouse movements

Inaccurate trials were excluded from the mouse movement analyses. In addition, trials with reaction times more than 2.5 standard deviations above the global mean (*M* = 7.80 ms; *SD* = 2.40), and trials with reaction times less than the mean sentence duration, were excluded from the mouse movement analyses (7.91% of accurate trials). Trials with targets on the left vs. right side of the visual array were also combined by inverting the horizontal axis in the former (e.g., correct destinations are plotted to the right side of the visual array). As an online study, sampling rates varied across participants, and *zoo* (Zeileis & Grothendieck, 2005) was used to generate interpolated mouse trajectories. In order to depict the trajectories of participants' mouse movements across space, trials were normalised temporally by dividing each into 101 time points (e.g., see Spivey, Grosjean & Knoblich, 2005); average x and y coordinates are plotted in Figure 6. In order to depict the trajectories of participants' mouse movements across time, trials were aligned temporally; average y[1] coordinates are plotted in Figure 7. As expected, vertical mouse movements to the target at the bottom of the visual array were delayed (e.g., from "kiwi" to "rectangle") in the 2 vs. 1 referent context.

In order to assess garden path effects across conditions, vertical mouse movements

[1]Horizontal mouse movements (i.e., x coordinates) were also analysed. Horizontal mouse movements during the analysis window were significantly higher (i.e., reflecting less leftward movement to the target) in the 2 vs. 1 referent context, *Est*. = 5.62, *SE* = 2.66, $t = 2.12$, $p < .05$. In addition, the interaction of referent context and instruction type was significant, *Est*. = -10.73, *SE* = 5.31, $t = -2.02$, $p < .05$, revealing that the effect of referent context was more pronounced for ambiguous vs. unambiguous instructions. However, no other (e.g., individual differences) effects were significant.

between the offset of the second noun and mean sentence offset were analysed, capturing the (i.e., upward) movements of participants' trajectories to the incorrect destination above the target (e.g., see Figure 6). Mean y coordinates are plotted by condition in Figure 8A. Mean y coordinates were submitted to a by-participants mixed effects model with fixed effects of referent context, instruction type and vocabulary knowledge and random intercepts by participants. Results are reported in Table 4 and depicted in Figure 8B.

(Table 4 about here)

Vertical mouse movements were significantly lower (i.e., reflecting less upward movement to the incorrect destination above the target) in the 2 vs. 1 referent context, and they were significantly lower for unambiguous vs. ambiguous instructions. In addition, the significant interaction of referent context and instruction type revealed that the effect of referent context was more pronounced for ambiguous vs. unambiguous instructions. Finally, the significant interaction of referent context, instruction type and vocabulary knowledge revealed that with unambiguous instructions, the effect of referent context was more pronounced for participants with more vs. less vocabulary knowledge. In other words, with unambiguous instructions, the divergence between 1 referent (i.e., upward orientated unfilled symbols; see Figure 7B) vs. 2 referent (i.e., downward orientated unfilled symbols) contexts was significantly weaker for participants with less (i.e., on the left) vs. more (i.e., on the right) vocabulary knowledge. In contrast, with ambiguous instructions, the divergence between 1 referent (i.e., upward orientated filled symbols) vs. 2 referent (i.e., downward orientated filled symbols) contexts was consistent across participants. Similarly, a correlational analysis of the difference in vertical mouse movements between 1 vs. 2 referent contexts with unambiguous instructions revealed a significant correlation with vocabulary knowledge, $r(40) = 0.38$, $p < 0.05$. In contrast, a correlational analysis of the difference in vertical mouse movements between 1 vs. 2 referent contexts with ambiguous instructions revealed a non-significant correlation with vocabulary knowledge, $r(40) = -0.14$, $p = 0.36$.

Discussion

Conceptually replicating prior findings (e.g., Farmer, Cargill et al., 2007), participants in Experiment 2 made less garden path mouse movements to the incorrect destination in the 2 referent visual context. In addition, the divergence between 2 vs. 1 referent visual contexts with unambiguous instructions was significantly weaker for participants with less vs. more vocabulary knowledge. Consistent with Experiment 1, these results link constraint integration processes to vocabulary knowledge.

While eye movements have been widely used in the literature to investigate individual differences in constraint integration processes, mouse movements have seen less use (e.g., but see Anderson et al., 2011). These results support the sensitivity of mouse movements to individual differences in these processes, including among adults. Moreover, these results also suggest that this sensitivity extends to internet-mediated research, potentially supporting data collection from diverse samples with wide individual differences profiles (e.g., beyond undergraduates).

General discussion

Across two experiments, participants hearing syntactically ambiguous sentences like “Put the kiwi on the rectangle on the circle” made less garden path eye (Experiment 1) and mouse (Experiment 2) movements to the incorrect (e.g., rectangle) destination when the visual context included 2 vs 1 kiwi, conceptually replicating prior findings (e.g., Farmer, Cargill et al., 2007; Tanenhaus et al., 1995). In addition, the divergence between 2 vs. 1 referent visual contexts was significantly weaker for participants with less vs. more vocabulary knowledge; this interaction spanned both ambiguous and unambiguous instructions in Experiment 1 and centred on unambiguous instructions in Experiment 2. These results provide novel insight into the integration of visual and linguistic constraints during sentence comprehension: building on research with children (e.g., Anderson et al., 2011; Ferreira et al., 2013; Snedeker & Trueswell, 2004; Trueswell et al., 1999; Woodard et al., 2016) and other special populations (e.g., Rabagliati et al., 2019), these results reveal a continuum of individual comprehender differences linked to vocabulary knowledge in adults.

The results of Experiments 1 and 2 suggest that capacities related to the integration of visual and linguistic constraints exist along a continuum. In contrast, a number of prior studies have emphasised qualitative distinctions; for example, Trueswell et al. (1999) found that adults but not children were constrained by the visual context. One explanation of Trueswell et al.'s findings is that adults and children differ because they rely on qualitatively distinct processes. However, the current results revealed that even adults varied in the degree to which they were constrained by the visual context, with less skilled adults (i.e., as reflected in their vocabulary knowledge) showing more child like (non)influences of visual constraints. In related research, Novick et al. (2008) also found that the degree to which adults were constrained by the visual context was related to their performance when reading syntactically ambiguous sentences. Taken together, these findings are consistent with the hypothesis that comprehenders' sensitivity to the visual context develops continuously over the course of language acquisition, such that children begin at the non-integration extreme of the continuum and develop greater and greater sensitivity to these constraints across language learning (e.g., showing adult like eye movement patterns by eight years of age; Weighall, 2008). At the same time, the current results extend prior research by suggesting that the endpoints of this development may be variable, such that even adults differ in their sensitivities. Finally, these results also suggest that adults' varying (in)sensitivity is linked to both (i.e., participant) internal (e.g., vocabulary knowledge) and external (e.g., visual preview and complexity; Ferreira et al., 2013) factors.

The results of Experiments 1 and 2 also suggest that individual differences in the integration of visual and linguistic constraints may be linked to vocabulary knowledge. Building on Anderson et al.'s (2011) findings with children, these results suggest that vocabulary knowledge is also a predictor of adults' eye and mouse movement patterns. The reading literature has focused considerable attention on the relationship between reading comprehension and vocabulary knowledge. According to the Lexical Quality Hypothesis (e.g., Perfetti, 2007), skilled reading comprehension depends on high quality lexical representations, which are characterised by detailed representations of form and rich representations of meaning. Prior individual differences research provides compelling support for this hypothesis; for example, Freed, Hamilton and Long (2017) found that their comprehensive battery of individual differences measures collectively accounted for 77% of the variance in reading comprehension, while vocabulary knowledge alone accounted for 72% (e.g., also see Braze, Katz, Magnuson, Mencl, Tabor, Van Dyke, Gong, Johns & Shankweiler, 2016; Braze, Tabor, Shankweiler & Mencl, 2007). Paralleling the visual domain (i.e., reading), the current results suggest that poorer lexical representations may also impair sentence comprehension in the auditory domain (i.e., speech). For example, Experiments 1 and 2 required comprehenders to map individual words from the speech stream onto individual objects in the visual context (e.g., the word “kiwi” onto the 1 or 2 kiwis in Figure 1). In addition, these experiments required comprehenders to track whether this lexical mapping necessitated modification (e.g., see 2 referent contexts). We conjecture that vocabulary knowledge provides an important constraint on these processes, such that the activation, mapping and coordination of this lexical information may be slower and less accurate in comprehenders with poorer lexical representations, who are thus less constrained by the visual context.

Interestingly, Experiments 1 and 2 revealed related but differing individual differences patterns. While Experiment 1 yielded a two way (i.e., Ref. x Vocab) interaction, such that effects spanned both ambiguous and unambiguous instructions, Experiment 2 yielded a three way (i.e., Ref. x Instr. x Vocab) interaction, such that effects centred on unambiguous instructions. These experiments differed in an important respect, which may explain this discrepancy: the visual arrays in the former were more complex (see also Ferreira et al., 2013), with three vs. two visual objects, and three vs. two source and destination locations. We conjecture that the influence of vocabulary knowledge may have been more pervasive in Experiment 1 because comprehenders with more vocabulary knowledge were better able to use the complex visual arrays to inhibit the incorrect destination irrespective of instruction type, underpinning the observed two way interaction. In contrast, the influence of vocabulary knowledge may have been less pervasive in Experiment 2 because all comprehenders (i.e., irrespective of vocabulary knowledge) were able to use the simple visual arrays to inhibit the incorrect destination. At the same time, comprehenders with more vocabulary knowledge may have been better able to integrate these visual constraints with the syntactic cues from the unambiguous instructions, underpinning the observed three way interaction.

The differing individual differences patterns observed in Experiments 1 and 2 may also reflect the differing (i.e., eye vs. mouse) measures in these experiments. A potential advantage of eye tracking over mouse tracking is its temporal sensitivity: typically, participants can move their eyes to a linguistically relevant object faster than they can move their mouse. Thus, the former may be more closely time locked to, and revealing of, language processes. However, consistent with Anderson et al. (2011), Farmer, Anderson et al. (2007) and Farmer, Cargill et al. (2007), the current garden path effects were observed using mouse tracking during a similar analysis window as eye tracking. Additionally, the detection of a three way interaction, as observed via mouse tracking in Experiment 2, typically requires greater power than a two way interaction, as observed via eye tracking in Experiment 1, not vice versa. Thus, we conjecture that the discrepancy between Experiments 1 and 2 is unlikely to be due to the sensitivity of eye vs. mouse methods; rather, both reflect richly sensitive measures. Finally, a powerful advantage of mouse tracking over eye tracking is its suitability for internet-mediated research. The current results suggest that mouse tracking also provides a powerful internet-mediated tool for studying individual differences.

An important limitation of this research is that it focused on a single individual differences dimension. Thus, the relationship between constraint integration processes and vocabulary knowledge could be spurious, such that it may stem from the shared variance between vocabulary knowledge and other (i.e., unmeasured) linguistic and cognitive skills (e.g., see Van Dyke et al., 2014). On the one hand, the current study lays an important foundation, establishing that the integration of visual and linguistic constraints varies systematically across skilled adult comprehenders (e.g., contrasting with the similarities observed between native and non-native speakers; Contemori et al., 2018; Nakamura et al., 2020; Pozzan & Trueswell, 2016). On the other hand, an important direction for future research will be to explore the unique contributions of a wider array of individual differences dimensions to skilled adult comprehension (e.g., via a more comprehensive battery). We conjecture that vocabulary knowledge is unlikely to be the sole determinant of constraint integration processes. Rather, prior research suggests that non-linguistic skills such as executive function may also support constraint integration processes, as reflected in children's performance errors (e.g., Woodard et al., 2016; see also Qi et al., 2020). For example, the current experiment also required comprehenders to coordinate the activation and inhibition of referents (e.g., see the 2 referent context) and parses (e.g., destination vs. modification), which may rely on domain general executive function skills. Likewise, prior findings have directly linked the processing of syntactic ambiguities in adults to executive function (e.g., Engelhardt et al., 2017; Hsu & Novick, 2016; Novick, Hussey, Teubner-Rhodes, Harbison & Bunting, 2014).

On balance, while the effects of referent context in Experiments 1 and 2 are consistent with constraint based approaches (e.g., MacDonald et al., 1994; McRae et al., 1998; Trueswell & Tanenhaus, 1994), the observed interactions with vocabulary also emphasise the need for mechanistic explanations of individual differences in constraint integration processes. According to these approaches, the visual context reflects one of many constraints that compete to influence parsing. However, less clear from these approaches is why these constraints are integrated differently by different comprehenders, and the underpinning mechanisms. We conjecture that skilled adult comprehenders are unlikely to be blind to the visual context; rather, a more plausible interpretation is that comprehenders may differ in how efficiently they are able to integrate conflicting constraints and/or in how they weight conflicting constraints. Findings from Snedeker and Trueswell (2004) emphasise the former interpretation: later in processing, they found that five year olds did show sensitivity to the visual context, suggesting that they may simply require additional processing time. Assuming that poorer lexical representations (e.g., Perfetti, 2007) likewise impair constraint integration processes, one possibility is that less skilled comprehenders may also show greater sensitivity to the visual context if allowed additional processing time (e.g., perhaps the opposite manipulation to Ferreira et al., 2013). Alternatively, assuming that vocabulary knowledge provides a *proxy* for language experience (e.g., paralleling Anderson et al., 2011), another possibility is that these comprehenders may have less experience with visual constraints on parsing, thus placing less weight on this information (e.g., vs. linguistic constraints like the destination bias of “put”). Finally, the current correlational results also do not address the direction of causality; thus, another possibility is that greater sensitivity to visual constraints may also facilitate vocabulary learning. Taken together, the current results suggest that another important direction for future research will be to distinguish the roles of efficiency, weighting and learning, alongside implementing models that incorporate these mechanisms.

In conclusion, the current visual world results provide novel insight into the integration of visual and linguistic constraints during sentence comprehension. While prior individual differences research has focused on children (e.g., Anderson et al., 2011; Woodard et al., 2016), the current results suggest that important meaningful variation also exists among skilled adults comprehenders. In particular, these results support a link between vocabulary knowledge (and potentially language experience generally) and constraint integration processes and emphasise the need for mechanistic accounts of individual differences in these processes.

Acknowledgments

Radhika Chauhan and Lacey Carter are thanked for their assistance with data collection.

Declaration of interest statement

The authors report no conflict of interest.

References

Anderson, S.E., Farmer, T.A., Goldstein, M., Schwade, J., Spivey, M. (2011). Individual differences in measures of linguistic experience account for variability in the sentence processing skill of five-year-olds. In I. Arnon & E.V. Clark (Eds.), *Experience, variation and generalization: Learning a first language (Vol. 7)* (pp. 203-221). John Benjamins Publishing.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., Gong, T., Johns, C., & Shankweiler, D. P. (2016). Vocabulary does not complicate the Simple View of Reading. *Reading and Writing, 29*(3), 435-451.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities, 40*(3), 226-243.

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

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 687-696.

Coco, M. I., & Keller, F. (2015). The interaction of visual and linguistic saliency during syntactic ambiguity resolution. *Quarterly Journal of Experimental Psychology, 68*(1), 46-74.

Contemori, C., Pozzan, L., Galinsky, P., & Dussias, P. E. (2020). When actions and looks don't line up: The contribution of referential and prosodic information in the processing of PP ambiguities in child-L2 speakers of English. *Linguistic Approaches to Bilingualism, 10*(5), 623-656.

Duñabeitia, J.A., Crepaldi, D., Meyer, A.S., New, B., Pliatsikas, C., Smolka, E., Brysbaert, M.

(2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology, 71*(4), 808-816.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test 4th edition (PPVT-4)*. Pearson.

Engelhardt, P. E., Nigg, J. T., & Ferreira, F. (2017). Executive function and intelligence in the resolution of temporary syntactic ambiguity: An individual differences investigation. *Quarterly Journal of Experimental Psychology, 70*(7), 1263-1281.

Farmer, T. A., Anderson, S. E., & Spivey, M. J. (2007). Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language, 57*(4), 570-595.

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., & Spivey, M. J. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science, 31*(5), 889-909.

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

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*(2), 178-210.

Freed, E. M., Hamilton, S. T., & Long, D. L. (2017). Comprehension in proficient readers: The nature of individual variation. *Journal of Memory and Language, 97*, 135-153.

Hintz, F., Meyer, A. S., & Huettig, F. (2017). Predictors of verb-mediated anticipatory eye movements in the visual world. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(9), 1352-1374.

Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological Science, 27*(4), 572-582.

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*(2), 151-171.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Knoeferle, P., & Guerra, E. (2016). Visually situated language comprehension. *Language and Linguistics Compass, 10*(2), 66-82.

Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., ... & Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica, 171*, 72-84.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676-703.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*(3), 283-312.

Nakamura, C., Arai, M., Hirose, Y., & Flynn, S. (2020). An extra cue is beneficial for native speakers but can be disruptive for second language learners: Integration of prosody and visual context in syntactic ambiguity resolution. *Frontiers in Psychology, 10*, 2835.

Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition and Neuroscience, 29*(2), 186-217.

Novick, J. M., Thompson-Schill, S. L., & Trueswell, J. C. (2008). Putting lexical constraints in context into the visual-world paradigm. *Cognition, 107*(3), 850-903.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Pozzan, L., & Trueswell, J. C. (2016). Second language processing and revision of garden-path sentences: a visual word study. *Bilingualism, 19*, 636-643.

Qi, Z., Love, J., Fisher, C., & Brown-Schmidt, S. (2020). Referential context and executive functioning influence children's resolution of syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(10), 1922-1947.

Rabagliati, H., Delaney-Busch, N., Snedeker, J., & Kuperberg, G. (2019). Spared bottom-up but impaired top-down interactive effects during naturalistic language processing in schizophrenia: evidence from the visual-world paradigm. *Psychological Medicine, 49*(8), 1335-1345.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics, 77*(3), 720-730.

Salverda, A. P., & Tanenhaus, M. K. (2017). The visual world paradigm. In A. M. De Groot & P. Hagoort (Eds.), *Research methods in psycholinguistics: A practical guide* (pp. 89-110). John Wiley & Sons.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*(3), 238-299.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences, 102*(29), 10393-10398.

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition, 131*(3), 373-403.

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*(5217), 1632-1634.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*(2), 89-134.

Trueswell, J., & Tanenhaus, M. (1994). Toward a lexicalist framework for constraint-based syntacticambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155-179). Lawrence Erlbaum.

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology, 99*(2), 75-95.

Woodard, K., Pozzan, L., & Trueswell, J. C. (2016). Taking your own path: Individual differences in executive function and language processing skills in child learners. *Journal of Experimental Child Psychology, 141*, 187-209.

Zeileis, A., & Grothendieck, G. (2005). zoo: S3 Infrastructure for regular and irregular time series. *Journal of Statistical Software, 14(6)*, 1-27.

Table 1.

Experiment 1: Mixed effects analysis of accuracy (Est., SE and 95% CI x 10^{-2})

Fixed Effects	***Est.***	***SE***	***95% CI***	***t***	***p***
Intercept	97.16	0.57	[96.04, 98.28]	171.72	< .001
Referent Context	-0.71	1.10	[-2.87, 1.45]	-0.65	0.52
Instruction Type	3.19	1.10	[1.03, 5.35]	2.90	< .01
Vocabulary Knowledge	1.10	0.57	[-0.02, 2.22]	1.92	0.05
Ref. x Instr.	-2.13	2.20	[-6.44, 2.18]	-0.97	0.33
Ref. x Vocab	-0.37	1.11	[-2.55, 1.81]	-0.33	0.74
Instr. x Vocab	-1.49	1.11	[-3.67, 0.69]	-1.35	0.18
Ref. x Instr. x Vocab	-0.34	2.22	[-4.69, 4.01]	-0.15	0.88

Table 2.

Experiment 1: Mixed effects analysis of (i.e., transformed) proportions of fixations to the incorrect destination (Est., SE and 95% CI x 10^{-2})

Fixed Effects	***Est.***	***SE***	***95% CI***	***t***	***p***
Intercept	58.59	3.79	[51.16, 66.01]	15.46	< .001
Referent Context	-36.46	2.97	[-42.27, -30.64]	-12.28	< .001
Instruction Type	-14.43	2.97	[-20.24, -8.61]	-4.86	< .001
Vocabulary Knowledge	-1.03	3.83	[-8.54, 6.48]	-0.27	0.79
Ref. x Instr.	19.03	5.94	[7.40, 30.67]	3.21	< .01
Ref. x Vocab	-7.10	3.00	[-12.98, -1.22]	-2.37	< .05
Instr. x Vocab	-1.51	3.00	[-7.39, 4.37]	-0.50	0.61
Ref. x Instr. x Vocab	-0.27	6.00	[-12.03, 11.49]	-0.05	0.96

Table 3.

Experiment 2: Mixed effects analysis of accuracy (Est., SE and 95% CI x 10^{-2})

Fixed Effects	***Est.***	***SE***	***95% CI***	***t***	***p***
Intercept	91.90	1.50	[88.96 ,94.84]	61.37	< .001
Referent Context	0.95	1.65	[-2.28 ,4.18]	0.58	0.56
Instruction Type	-0.48	1.65	[-3.71 ,2.75]	-0.29	0.77
Vocabulary Knowledge	3.40	1.52	[0.42 ,6.38]	2.24	< .05
Ref. x Instr.	1.27	3.29	[-5.18 ,7.72]	0.39	0.70
Ref. x Vocab	-0.94	1.67	[-4.21 ,2.33]	-0.56	0.58
Instr. x Vocab	-0.35	1.67	[-3.62 ,2.92]	-0.21	0.83
Ref. x Instr. x Vocab	4.92	3.33	[-1.61 ,11.45]	1.48	0.14

Table 4.

Experiment 2: Mixed effects analysis of vertical mouse movements (Est., SE and 95% CI x 10^{-2})

Fixed Effects	***Est.***	***SE***	***95% CI***	***t***	***p***
Intercept	-31.31	2.34	[-35.90, -26.72]	-13.36	< .001
Referent Context	-13.09	2.73	[-18.44, -7.74]	-4.79	< .001
Instruction Type	-17.76	2.73	[-23.11, -12.41]	-6.50	< .001
Vocabulary Knowledge	3.62	2.37	[-1.03, 8.27]	1.52	0.13
Ref. x Instr.	24.58	5.46	[13.88, 35.28]	4.50	< .001
Ref. x Vocab	-3.40	2.77	[-8.83, 2.03]	-1.23	0.22
Instr. x Vocab	-2.01	2.77	[-7.44, 3.42]	-0.73	0.47
Ref. x Instr. x Vocab	-11.74	5.53	[-22.58, -0.90]	-2.12	< .05

Figure 1. Experiment 1: Visual arrays for the example instruction, “Put the kiwi (that’s) on the rectangle on the circle”. While the 1 referent context (left) included one kiwi, the 2 referent context (right) included two kiwis. Participants were tasked with using the mouse to move one visual object on each trial from a pink source to gray destination location.

Figure 2. Experiment 1: Average (shaded bands show 95% CIs) proportions of fixations to target (e.g., kiwi on pink rectangle) and distractor (e.g., shirt on pink triangle) referents and correct (e.g., gray circle) and incorrect (e.g., gray rectangle) destinations between sentence onset and offset (+600 ms) in the 1 referent context. Participants heard ambiguous (A) vs. unambiguous (B) instructions like, “Put the kiwi (that’s) on the rectangle on the circle”.

Figure 3. Experiment 2: Average (shaded bands show 95% CIs) proportions of fixations to target and competitor referents and correct and incorrect destinations between sentence onset and offset (+600 ms) in the 2 referent context with ambiguous (A) vs. unambiguous (B) instructions.

Figure 4. Experiment 1: Average (*SE*) proportions of fixations by (referent context x instruction type) condition to the incorrect destination between the offset of the second noun and sentence offset (e.g., “…on the circle”) (A) and model fits from the mixed effects analysis of (i.e., transformed) incorrect destination fixations (B).

Figure 5. Experiment 2: Visual arrays for the example instruction, “Put the kiwi (that’s) on the rectangle on the circle” in the 1 (left) vs. 2 (right) referent context. In contrast to Experiment 1, visual arrays included only two visual objects (i.e., and source and destination locations) to constrain participants’ patterns of mouse movements and enable these to be aggregated across trials in the analyses.

Figure 6. Experiment 2: Temporally normalised average mouse trajectories across the visual

array in the 1 vs. 2 referent context. Participants heard ambiguous (A) vs. unambiguous (B) instructions like, “Put the kiwi (that’s) on the rectangle on the circle”. The centre of the visual array is (0, 0). The target (e.g., kiwi on pink rectangle) is plotted at the bottom left of the visual array (i.e., -1, -1), the correct destination (e.g., gray circle) is plotted at the top right of the visual array (i.e., 1, 1) and the incorrect destination (e.g., gray rectangle) is plotted at the top left of the visual array (i.e., -1, 1).

Figure 7. Experiment 2: Average (shaded bands show 95% CIs) vertical mouse movements (i.e., y coordinates) between sentence onset and offset (+600 ms) in the 1 vs. 2 referent context with ambiguous (A) vs. unambiguous (B) instructions.

Figure 8. Experiment 2: Average (*SE*) y coordinates by (referent context x instruction type) condition between the offset of the second noun and mean sentence offset (e.g., “…on the circle”) (A) and model fits from the mixed effects analysis of y coordinates (B).

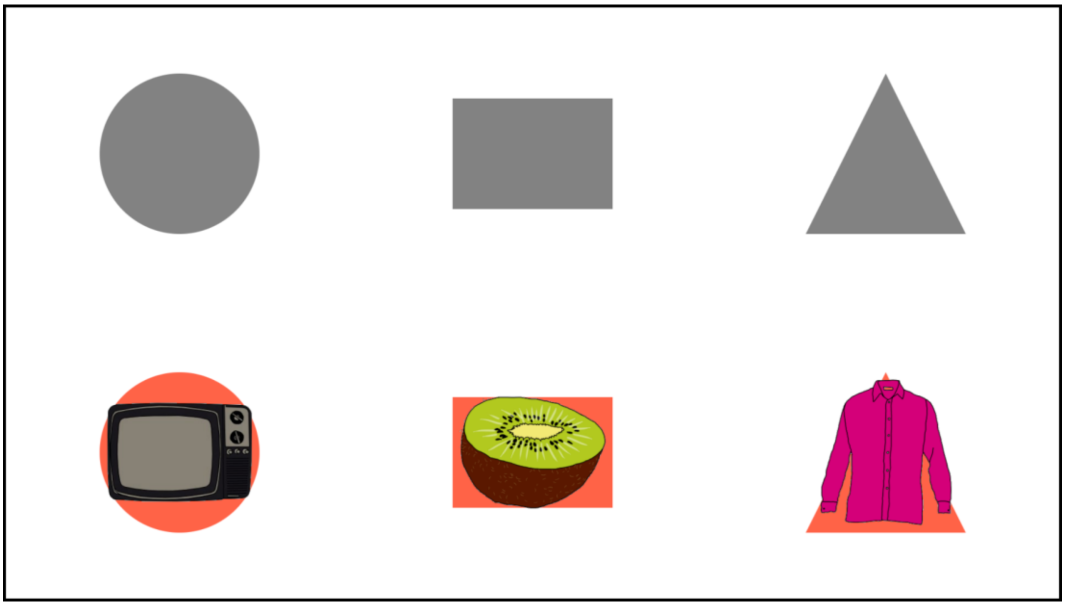

Figure 1A.

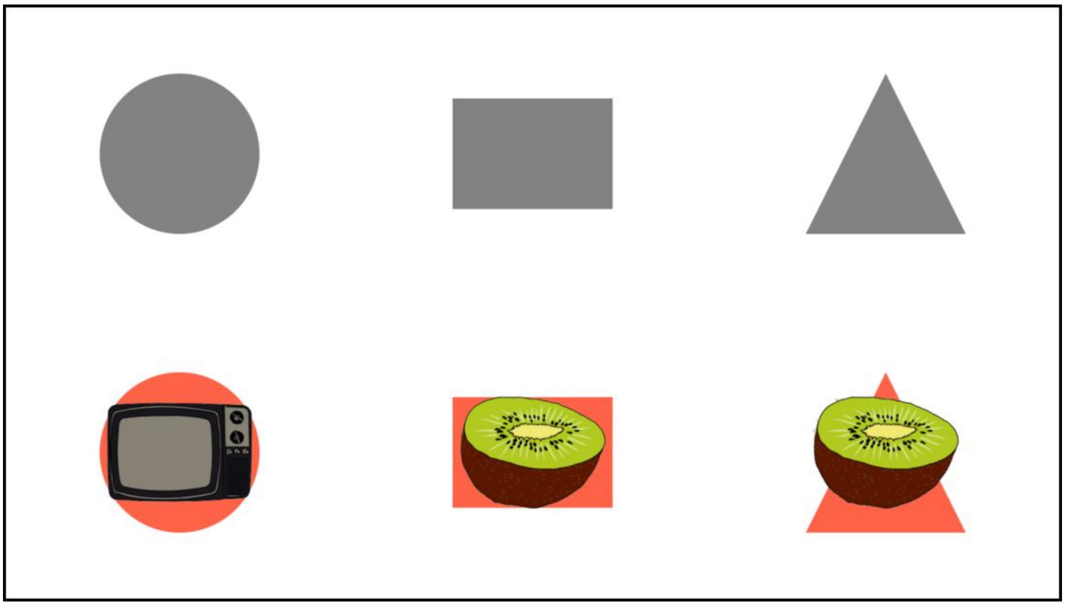

Figure 1B.

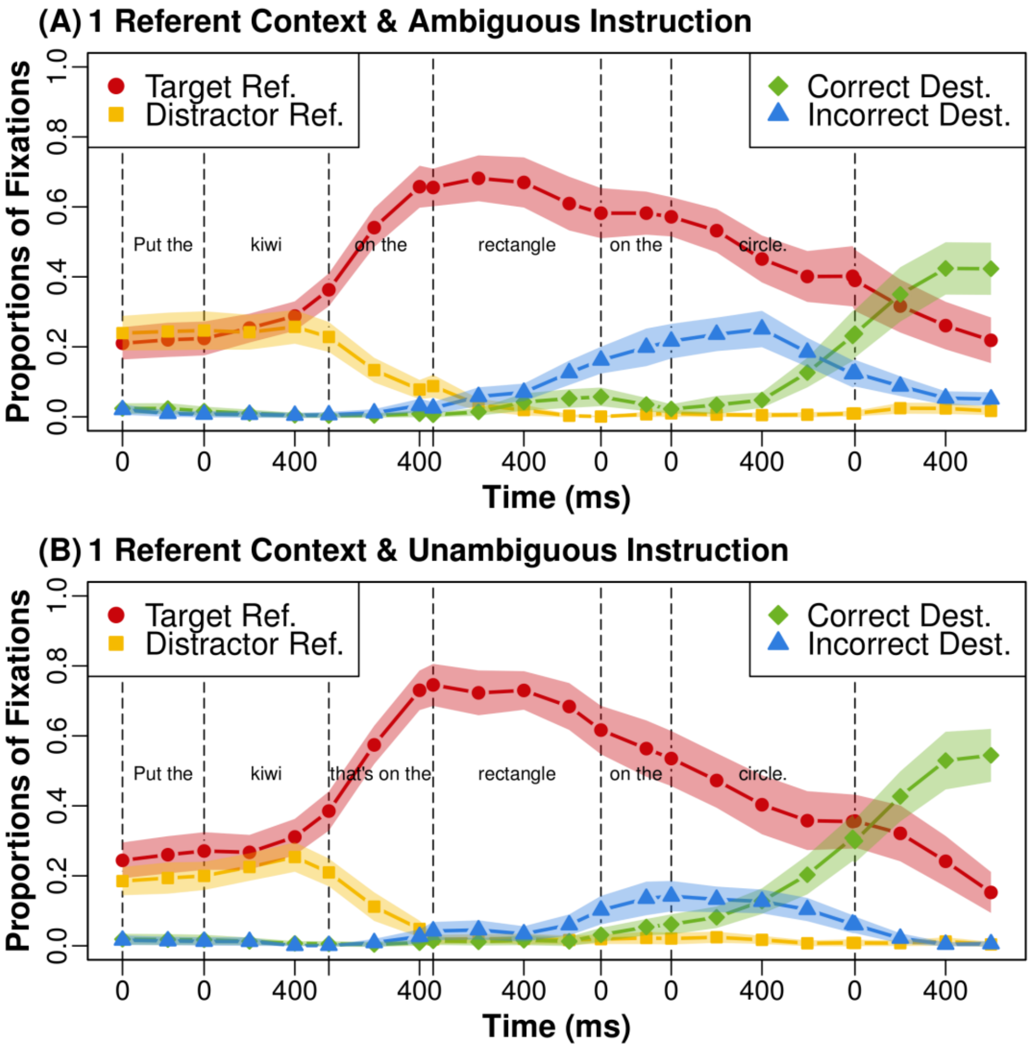
(A) 1 Referent Context & Ambiguous Instruction
Target Ref.
Distractor Ref.
Correct Dest.
Incorrect Dest.
Put the
kiwi
on the
rectangle
on the
circle.
Proportions of Fixations
Time (ms)
(B) 1 Referent Context & Unambiguous Instruction
Target Ref.
Distractor Ref.
Correct Dest.
Incorrect Dest.
Put the
kiwi
that's on the
rectangle
on the
circle.
Proportions of Fixations
Time (ms)

Figure 2.

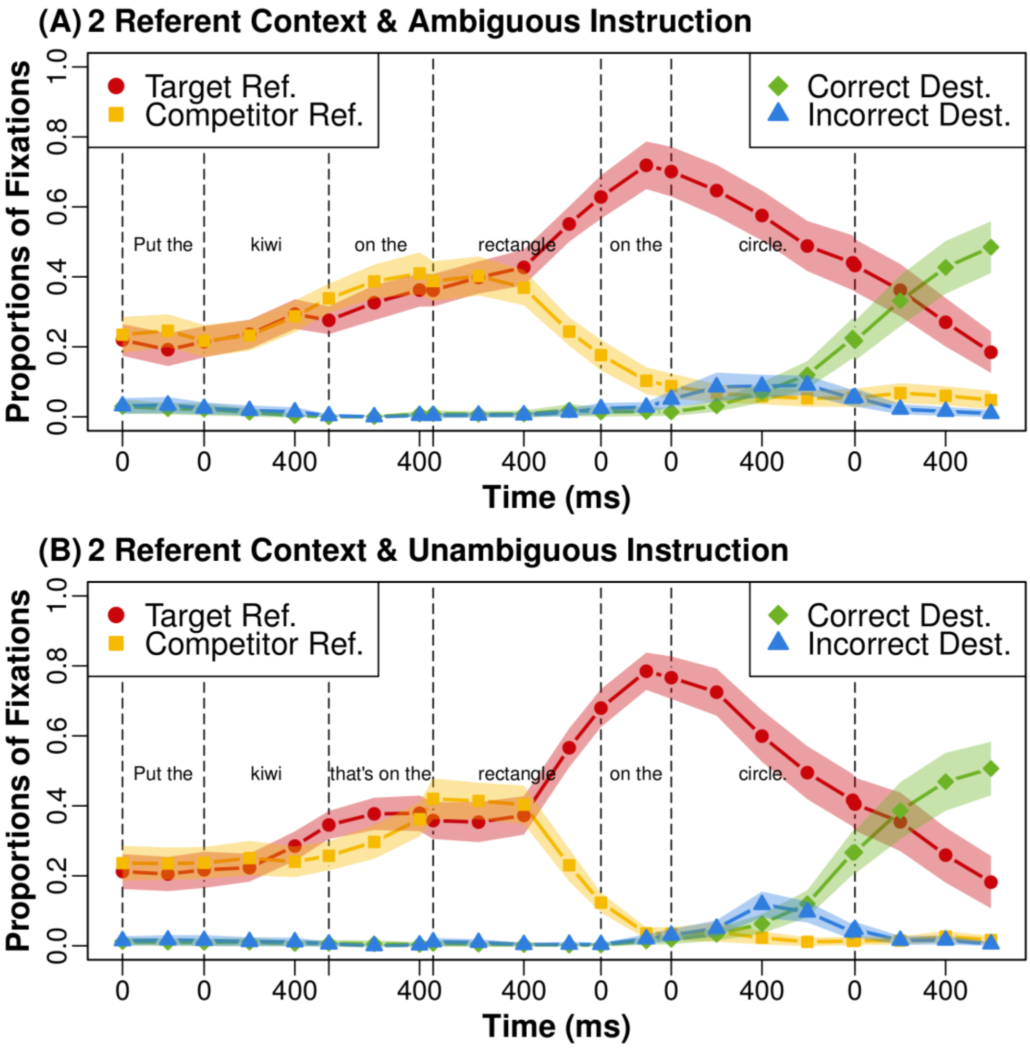
(A) 2 Referent Context & Ambiguous Instruction
Proportions of Fixations
Target Ref.
Competitor Ref.
Correct Dest.
Incorrect Dest.
Put the
kiwi
on the
rectangle
on the
circle.
Time (ms)
(B) 2 Referent Context & Unambiguous Instruction
Proportions of Fixations
Target Ref.
Competitor Ref.
Correct Dest.
Incorrect Dest.
Put the
kiwi
that's on the
rectangle
on the
circle.
Time (ms)

Figure 3.

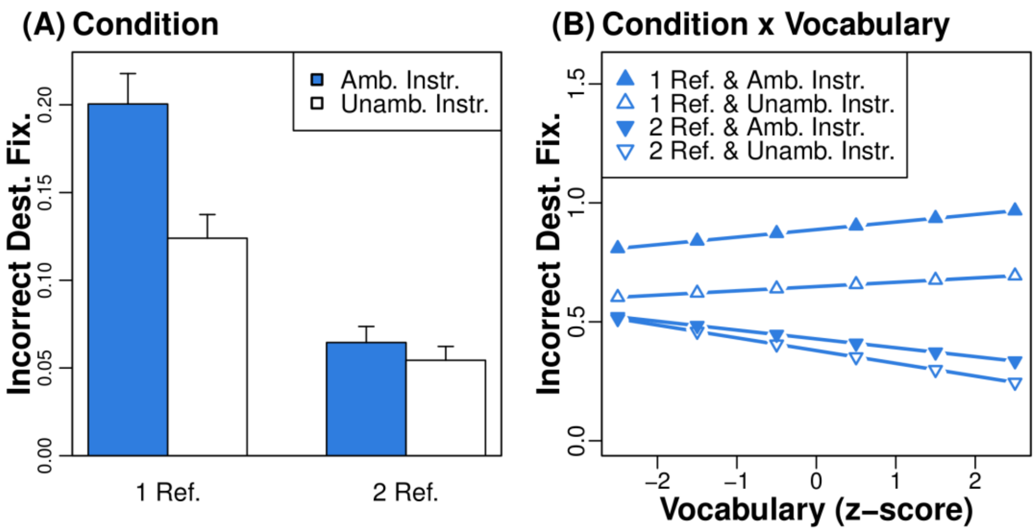
(A) Condition
Amb. Instr.
Unamb. Instr.
Incorrect Dest. Fix.
0.00
0.05
0.10
0.15
0.20
1 Ref.
2 Ref.
(B) Condition x Vocabulary
1 Ref. & Amb. Instr.
1 Ref. & Unamb. Instr.
2 Ref. & Amb. Instr.
2 Ref. & Unamb. Instr.
Incorrect Dest. Fix.
0.0
0.5
1.0
1.5
−2
−1
0
1
2
Vocabulary (z−score)

Figure 4.

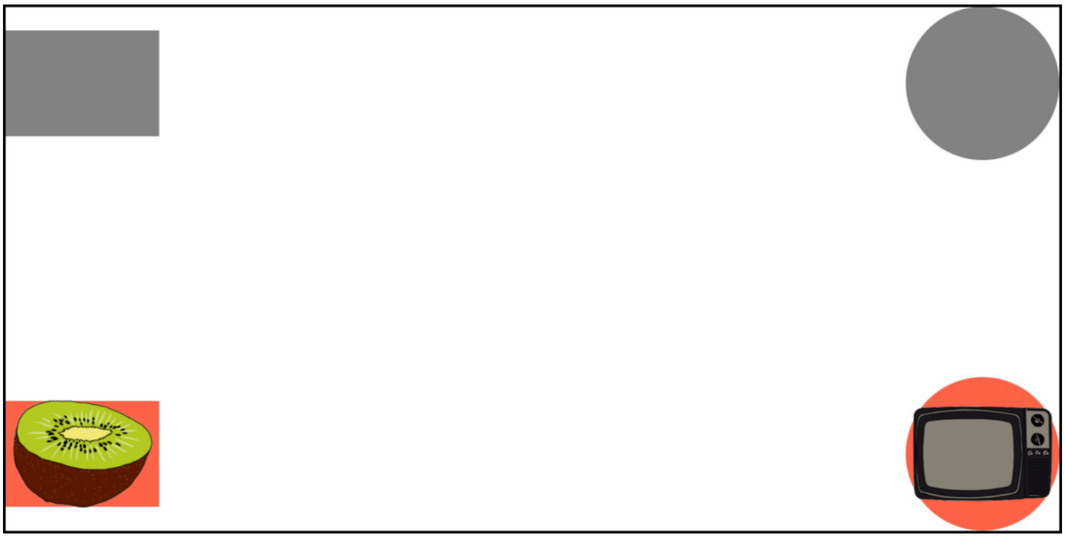

Figure 5a.

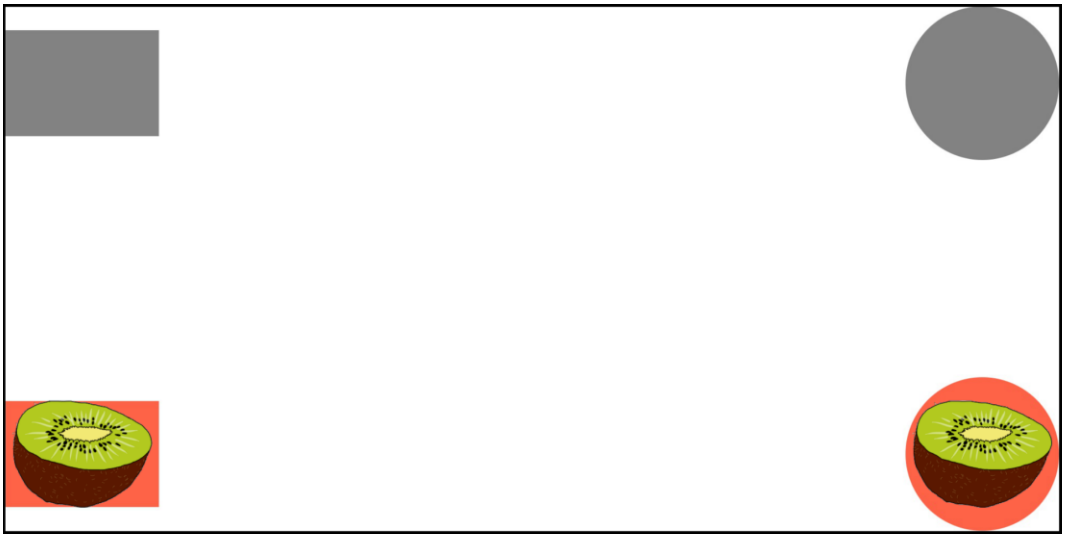

Figure 5b.

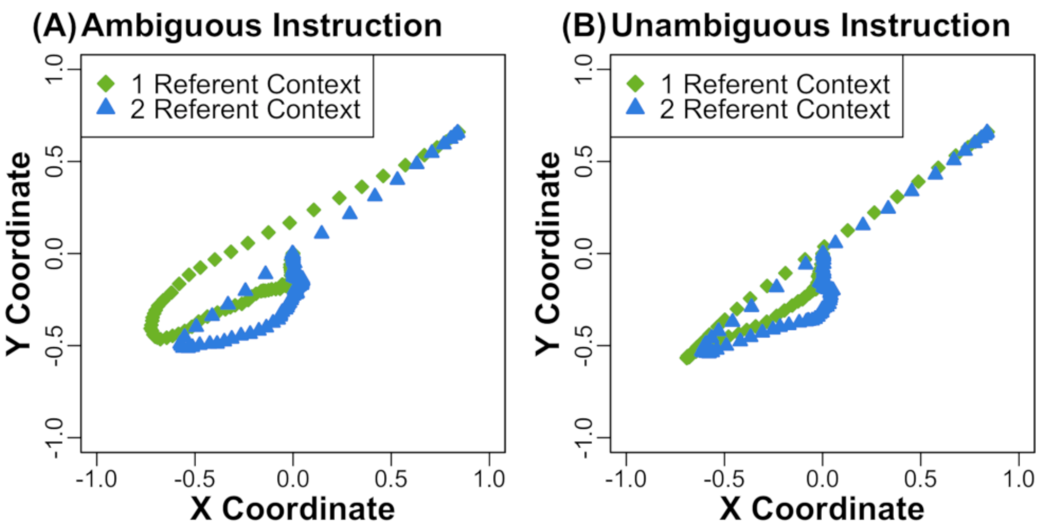
(A) Ambiguous Instruction
(B) Unambiguous Instruction
1 Referent Context
2 Referent Context
Y Coordinate
X Coordinate
-1.0
-0.5
0.0
0.5
1.0

Figure 6.

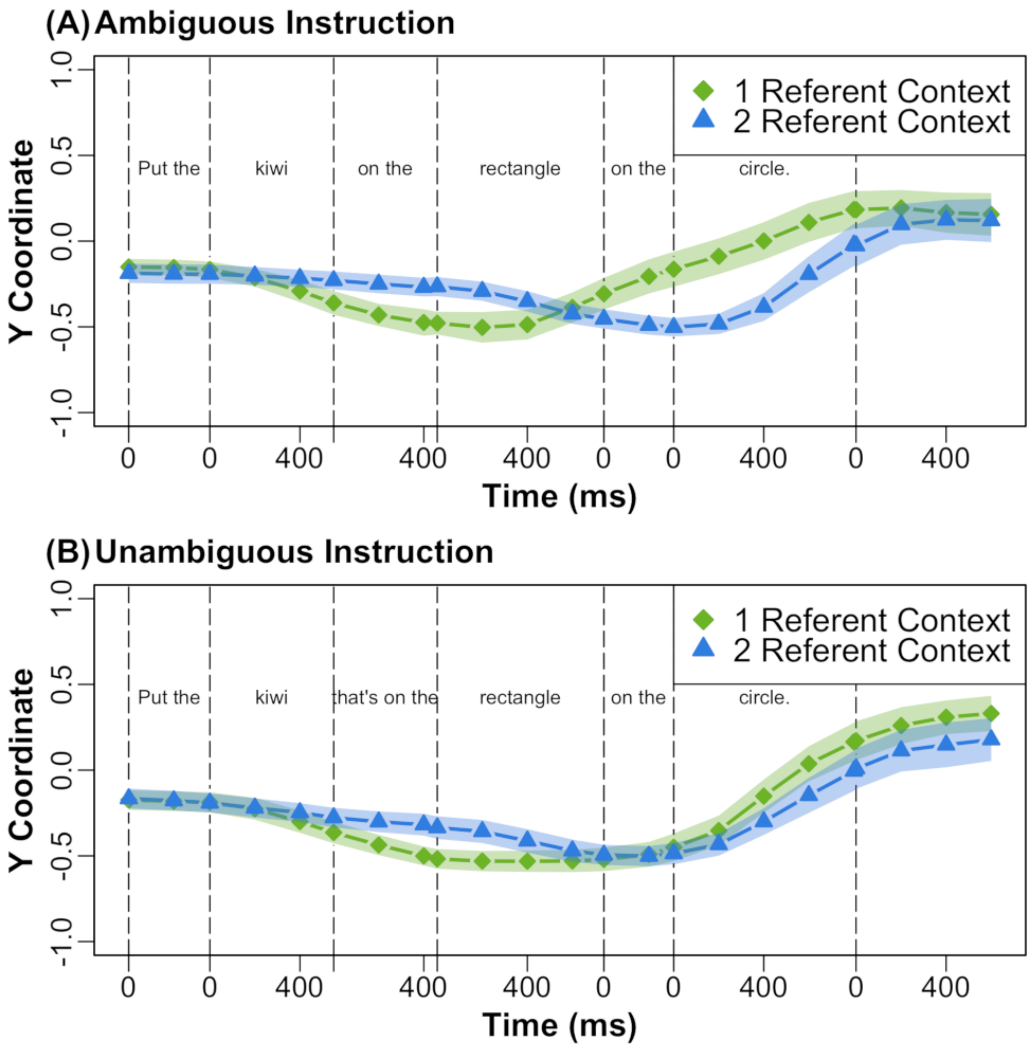
(A) Ambiguous Instruction
1 Referent Context
2 Referent Context
Put the
kiwi
on the
rectangle
on the
circle.
Y Coordinate
1.0
0.5
0.0
-0.5
-1.0
0
0
400
400
400
0
0
400
0
400
Time (ms)
(B) Unambiguous Instruction
1 Referent Context
2 Referent Context
Put the
kiwi
that's on the
rectangle
on the
circle.
Y Coordinate
1.0
0.5
0.0
-0.5
-1.0
0
0
400
400
400
0
0
400
0
400
Time (ms)

Figure 7.

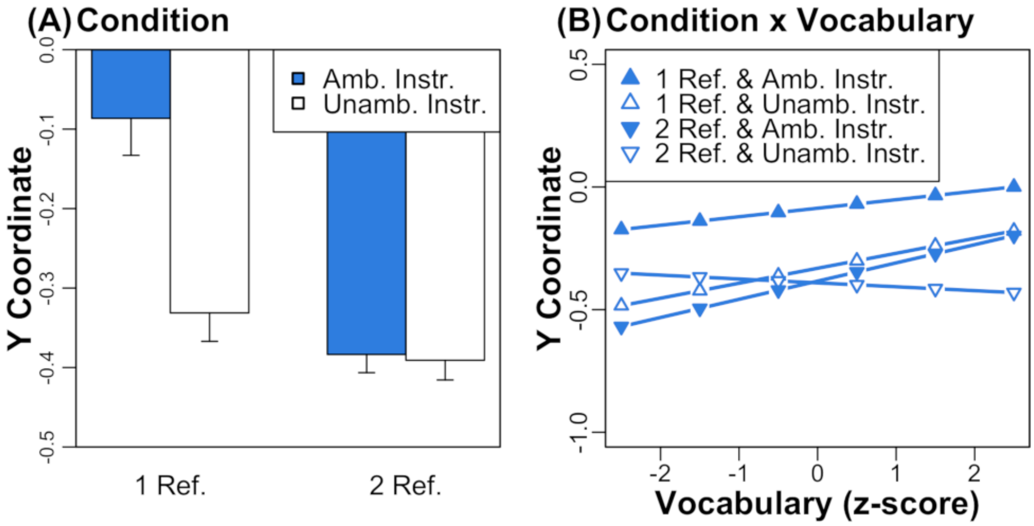
(A) Condition
Amb. Instr.
Unamb. Instr.
Y Coordinate
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
1 Ref.
2 Ref.
(B) Condition x Vocabulary
1 Ref. & Amb. Instr.
1 Ref. & Unamb. Instr.
2 Ref. & Amb. Instr.
2 Ref. & Unamb. Instr.
Y Coordinate
0.5
0.0
-0.5
-1.0
-2
-1
0
1
2
Vocabulary (z-score)

Figure 8.