

Eye Gaze as a Measure of the Effects of Syntactic Complexity and Agency Distance on the
Speed of Spoken Language Comprehension: A New Paradigm

Master's Thesis

Presented to

The Faculty of the Graduate School of Arts and Sciences
Brandeis University
Department of Psychology
Dr. Arthur Wingfield, Advisor

In Partial Fulfillment
of the Requirements for the Degree

Master of Arts
in
Psychology

by
Victoria Sorrentino

May 2017

Copyright by
Victoria Sorrentino

© 2017

ABSTRACT

Eye Gaze as a Measure of the Effects of Syntactic Complexity and Agency Distance on the Speed of Spoken Language Comprehension: A New Paradigm

A thesis presented to the Department of Psychology

Graduate School of Arts and Sciences
Brandeis University
Waltham, Massachusetts

By Victoria Sorrentino

In 1999, Caplan and Waters suggested that basic sentence comprehension is an on-line process that does not draw upon working memory resources. While most spoken language comprehension tasks use off-line measures, the present study additionally examined eye-gaze in a speech comprehension task that manipulated syntactic complexity and agency distance. Participants listened to spoken sentences and were asked to click on the “doer” of the action in each sentence from two choices presented on a computer screen in front of them. Among eighteen younger adults, it was found that the times required to click on the correct answers (overt response times or ORTs) were more affected by the manipulations than the times required to fixate on the correct answers (eye-fixation times or EFTs), and that working memory span had a stronger relationship with on-line responses than off-line responses. These results provide partial support for Caplan and Waters’ model; working memory span did not predict EFTs in more syntactically complex sentences, but effects of working memory span were seen on EFTs based on agency distance, suggesting that working memory resources were in fact required for the comprehension task.

Table of Contents

Abstract	iii
List of Tables	v
List of Figures	vi
Introduction	1
Methods	6
Participants	
Vision and Hearing Screening	
Cognitive Testing	
Speech Stimuli	
Visual Display	
Procedure	
Results	11
Accuracy	
Response Time	
Working Memory Span and Inhibitory Control Ability	
Discussion	16
Conclusion	21
References	22
Tables	25
Figures	32

List of Tables

Table 1. Sample Sentences of Each Type.	25
Table 2. Means and Standard Deviations for Accuracy by Sentence Type	26
Table 3. Means and Standard Deviations for Overt Response Times by Sentence Type in Seconds.	27
Table 4. Means and Standard Deviations for Eye-Fixation Times by Sentence Type in Seconds.	28
Table 5. Zero-Order and Partial Correlations for (a) Accuracy, (b) Working Memory Span, and (c) Inhibitory Control Ability	29
Table 6. Zero-Order and Partial Correlations for (a) Overt Response Time, (b) Working Memory Span, and (c) Inhibitory Control Ability.	30
Table 7. Zero-Order and Partial Correlations for (a) Eye-Fixation Time, (b) Working Memory Span, and (c) Inhibitory Control Ability.	31

List of Figures

Figure 1. Average proportion of time spent fixating to the agent or “doer” in each sentence type. 32

Figure 2. Accuracy by complexity and agency distance. 33

Figure 3. Overt response time and eye-fixation time by complexity and agency distance 34

INTRODUCTION

Language is a critical aspect of communication among humans. Constantly, people are presented with questions, conversations, and information through spoken language, and the comprehension of these messages plays a key role in everyday life. Simply hearing speech can be difficult with fast speech rates and background noise, while understanding the syntax, giving it meaning, and making use of that information is an additional feat (Helfer & Staub, 2014; Wingfield, Peelle, & Grossman, 2003).

Baddeley and Hitch (1994) divided working memory into three components: the central executive and its two slave systems, the visuospatial sketchpad and the phonological loop. Caplan and Waters (1999) posed a further specialization of verbal working memory into subsystems, breaking down sentence comprehension into interpretive and post-interpretive processing, where interpretive processing involves understanding the basic meaning of a sentence (syntactic assignment and thematic roles) and post-interpretive processing involves using that meaning to reason, plan, and accomplish other tasks. They posed a separate sentence interpretation resource (SSIR) theory, suggesting that a specific part of the verbal working memory system is used for interpretive aspects of sentence comprehension. This contrasts the widely accepted single resource (SR) theory, which posits that sentence comprehension and working memory tasks draw on the same resources (e.g. Just & Carpenter, 1992). Therefore, according to Caplan and Waters (1999), working memory performance does not predict language processing efficiency. They argue that on-line interpretive processing does not use any measurable working memory resources, thus explaining why reduced working memory

capacities, as in older adults, produce no impairments in simple sentence comprehension (Caplan & Waters, 1999; Waters & Caplan, 2001; see also Caplan & Waters, 2013; Caplan, Michaud, & Hufford, 2013; Caplan, Dede, Waters, Michaud, & Tripodis, 2011). However, as the working memory demands of sentences increase or the sentences become more complex, comprehension differences according to an on-line measure may in fact appear. The present study examined these manipulations in a spoken language comprehension task using both off-line and on-line measures.

A common procedure for studying spoken language comprehension involves answering a question after a sentence has been heard. However, these studies generally entail verbal or manual responses, such as a button press, to indicate the selected answer, thus requiring verbal or motor skills along with the comprehension task. In contrast, on-line measures of reaction time are “interpreted as reflecting the demands of processing of words and phrases as they are perceived and integrated into the sentence” (Waters & Caplan, 2001, p. 129). While older adults have been found to be slower to understand speech than younger adults using off-line methods (Wingfield, Peelle, & Grossman, 2003; Tun, Benichov, & Wingfield, 2010; Yoon et. al., 2015), comparable speech comprehension for these age groups has been found through on-line measures such as eye-tracking (Ayasse, Lash, & Wingfield, 2016), brain imaging (Wingfield & Grossman, 2006), and the auditory moving windows (AMW) task (Waters & Caplan, 2001), at least for sentences with simple syntax.

The visual world paradigm links the understanding of spoken language with vision (Alloppenna, Magnuson, & Tanenhaus, 1998), and past studies have found that eye movements occur in close proximity to relevant parts of a spoken story (Tanenhaus, Magnuson, Dahan, & Chambers, 2000; Huettig, Rommers, & Meyer, 2011). People tend to look at the items they hear,

or related items, as they hear them; therefore, eye tracking, an on-line measure, can be used to measure when people comprehend what they heard (Cooper, 1974). This takes advantage of the fact that eye-fixations are very rapid, occurring long before an overt response can be made. Further, as overt responses can add an additional task of selecting an answer and are inherently off-line indices of when a sentence has been understood, eye tracking can provide an alternative on-line, continuous measure of comprehension as the sentence is being presented (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Wendt, Kollmeier, & Brand, 2015; Wendt, Brand, & Kollmeier, 2014; Hadar, Skrzypek, Wingfield, & Ben-David, 2016).

A recent study by Ayasse, Lash, and Wingfield (2016) found that while off-line overt responses indicated seemingly slower language comprehension for older adults, older and younger adults were not significantly different when indexed by eye-fixation times (EFTs). In this study, participants saw four pictures on a screen, one of which was referred to in a spoken sentence. Participants were asked to use a mouse to click on the picture corresponding to an object referenced by the sentence, which was always the last word of the sentence. For example, participants heard a sentence such as, “She opened the door with a...” with a key being one of the four pictured objects. Both overt response times (ORTs) and EFTs were measured from the earliest point that a listener would be likely to know the identity of the referenced object, as indicated by a preliminary study using a cloze procedure. As expected, the time to select the referenced word with the mouse-click (ORT) was longer for older adults. On the other hand, the EFTs were similar for both age groups. These results imply that the difference in ORTs between older adults and younger adults was influenced by the time needed for older adults to make post-comprehension decisions and to physically select their answer with a mouse, not necessarily to comprehend the sentences. This study’s eye tracking results suggest that age differences in the

speed of sentence comprehension may be minimal for sentences with simple, non-demanding syntax. These findings are consistent with Caplan and Waters' distinction between interpretive and post-interpretive processing (1999), in that on-line interpretive processing of speech was not affected by the working memory differences associated with aging (e.g. Salthouse & Babcock, 1991).

The goal of my thesis is to use the on-line measure of eye-tracking while listening to sentences to test Caplan and Waters' postulated processing distinction and the role of working memory in affecting each of the stages of comprehension. It has been found that object-relative (OR) sentences (i.e. Boys that girls confuse are mean) are more difficult to understand than subject-relative (SR) sentences (i.e. Girls that confuse boys are mean) because not only are OR sentences less common in everyday language, they also violate the listener's expectations and require a reanalysis of the sentences (Wingfield et al., 2003; DeCaro, Peelle, Grossman, & Wingfield, 2016; Gibson, Bergen, & Piantadosi, 2013; Wendt et al., 2015; see also Novick, Trueswell, & Thompson-Schill, 2005). Syntactically complex sentences are associated with decreased comprehension, increased processing time, and increased effort (Wendt, Dau, & Hjortkjaer, 2016). Hearing a longer sentence also increases working memory load, as the listener must remember the elements of the sentence for a longer period of time (Just & Carpenter, 1992). However, according to Caplan and Waters' model, and in contrast to a number of others, OR sentences do not tax working memory any more than SR sentences (Ferreira et al., 1996; Cooke et al., 2002; Just & Carpenter, 1992; Wingfield et al., 2006; Zurif et al., 1995).

I tested the hypothesis that increasing the distance between the agent of an action and the action being described in a sentence (agency distance) would show differences in on-line eye-

gaze times due to a presumed increase in working memory demands. This is expected to tax working memory resources by increasing the amount of time required to hold the agent in memory. The present study manipulated syntactic complexity and agency distance to see if the speed of sentence comprehension, as measured by times to eye fixation on a referenced word, was reduced in more complex sentences that were thought to tax working memory resources. It was hypothesized that differences in eye gaze would be found based on agency distance, since this manipulated working memory demands, but there would be minimal differences based on complexity because increasing complexity does not in itself tax working memory. Specifically, long separation sentences were expected to produce greater EFTs than short separation sentences, and differentially so for OR sentences. Support for this hypothesis would provide evidence against Caplan and Waters, indicating that comprehension of complex sentences does in fact require working memory resources. As has been previously found, it is also expected that OR long separation sentences will result in slower overt responses than any other sentence type.

METHODS

Participants

Twenty-three native English-speaking young adults (YAs) participated in the study. Participants were recruited through fliers and online listservs, and all participants were compensated for their time. Due to the eye-tracker's difficulty in tracking some participant's pupils or clear misunderstanding of the task, five YAs were excluded from all analyses, leaving eighteen YAs (12 female, mean age = 21.67 years).

Vision and Hearing Screening

A vision screening was conducted using a Snellen eye chart (Hetherington, 1954) at 20 feet and the Jaeger close vision eye test (Holladay, 2004) at 12 inches. All participants had corrected vision of 20/50 or better. Hearing was tested for pure-tone acuity and speech reception thresholds (SRT) using a Grason-Stadler AudioStar Pro clinical audiometer (Grason-Stadler, Inc., Madison, WI) played through Eartone 3A insert earphones (E-A-R Auditory Systems, Aero Company, Indianapolis, IN). All participants had a pure-tone average (PTA) over .5, 1, 2, and 4 kHz < 25 dB HL, thus meeting a criterion of clinically normal hearing for speech (Katz, 2002).

Cognitive Testing

Three ancillary tests were administered to participants: (1) Vocabulary. The Shipley vocabulary test (Zachary, 1986) was used as a screening tool to ensure that, beyond self-report, participants had a general understanding of the English language. (2) Working Memory (WM) Span. Participants had verbal working memory tested using the Reading Span task based on Daneman and Carpenter (1980). Reading Span is a working memory task that asks participants

to determine if sentences are true or false while also remembering the last word of each sentence for later recall. The number of sentences in each block ranges from one through a maximum of five, increasing throughout the task every three trials. Participants were considered correct if the correct words were recalled in the order they were presented. If participants were incorrect on two of the three trials of a block at the three-word level or higher, the task was over. Their score was the sum total of the number of correct trials, with a maximum possible score of 15. (3)

Inhibitory Control Ability. Inhibitory control ability was tested with the Eriksen Flanker task (Eriksen & Eriksen, 1974). The Flanker task requires participants to ignore surrounding arrows and lines on a computer screen in order to quickly select the direction the center arrow is pointing. The irrelevant stimuli can either be arrows pointing in the same direction or the opposite direction as the center arrow, or horizontal lines without a direction. Participants were scored on their accuracy as well as reaction time, and the difference in reaction time between congruent and incongruent trials (conflict cost) was calculated and used for analyses. Reading Span and Flanker scores were used in post-hoc partial correlations, as WM span and inhibitory control ability are related constructs.

An additional task assessing participants' use of a computer mouse was also given. In this task, participants were asked to click on black dots as they appeared around the screen as quickly as they could. For the first twelve trials, they were told where the subsequent dot would be located, and for the remaining twelve trials, the dots appeared in random locations around the edge or center of the screen.

Participants provided written informed consent according to a protocol approved by the Brandeis University Committee for Protection of Human Subjects.

Speech Stimuli

The speech stimuli were previously used by DeCaro et al. (2016). These consisted of 864 spoken sentences, plus 52 filler sentences, recorded by a female native English speaker with natural intonation at an average speaking rate of 150 words per minute using Sound Studio v2.2.4 (Macromedia, Inc., San Francisco, CA, USA) that digitized (16-bit) at a sampling rate of 44.1 kHz. Recordings were equalized within and across sentence types for root-mean-square (RMS) intensity using MATLAB (MathWorks, Natick, MA, USA).

The experimental sentences varied in syntactic complexity (SR, OR), each containing a “doer” (agent) and receiver of an action. Some sentences contained an added four-word clause whose location in the sentence created either a short or long separation between the agent and the action. The long separation was intended to increase working memory demands by increasing the length of time the participant needed to keep the agent in mind before the action was heard. It should be noted that sentence length was controlled for, as short and long separation sentences were the same length, but the placement of the clause in these sentences was the key variable. As illustrated in Table 1, there were six total types of experimental sentences: base (no clause), short (clause did not additionally separate agent and action), and long (clause increased the separation between agent and action) in both SR and OR. Each version of the sentences was composed of the same key words, rearranged by sentence type.

Visual Display

As will be discussed in the Procedure section, two words appeared on a computer screen before each sentence was heard; one was the agent of the action, and the other was the recipient (or, for some fillers, the other group of people referenced in the sentence). The two words appeared at the same time, to the left and right of a small fixation circle in the center of the screen. Although the visual world paradigm originally used pictures of objects, a number of

subsequent studies have instead used printed words with good success (Salverda & Tanenhaus, 2010; McQueen & Viebahn, 2007; Huettig & McQueen, 2007).

Procedure

Participants were seated in front of a computer screen, with their head placed in an adjustable head rest 60cm from the eye tracker lens. Eye movements during each trial were recorded via an EyeLink 1000 Plus (SR Research Ltd., Ottawa, ON, Canada). The stimulus sentences were presented binaurally through the insert earphones at 20db higher than each participant's better ear SRT. An audibility check was performed prior to the practice trials to ensure that speech was intelligible for all participants. For the main task, participants were instructed to identify the agent or "doer" of the action in each sentence by clicking on one of two words presented on a screen, as described above. These two words were the people referenced in the accompanying sentence (i.e. "boys" and "girls").

In each trial, the words appeared on the screen three seconds before the centered fixation circle changed from white to black. After this color change, participants could use the computer mouse to click on the circle at any point, causing the sentence to play through the earphones. Participants were instructed to click on the word representing the agent of the action being described as soon as they thought they knew it. While the sentence played, participants were able to click on their response at any time, although the full sentence played regardless of when participants made their selection.

Participants listened to 144 experimental sentences and 52 filler sentences in a randomized order, with three optional breaks. Eight practice trials of the same format were done at the start of the experiment; these sentences were not used in the main task. There were six running orders; each core sentence appeared once per participant, in a different form in each

running order. Answer choices were presented in black writing on the left and right sides of a white computer screen. The entire task was run using MATLAB (MathWorks, Natick, MA, USA).

Accuracy, overt response time (ORT) and eye fixation time (EFT) were measured. ORT was defined as the time at which the participant used the mouse to click on the correct word on the screen, and the first mouse click after the start of the sentence was also the measure of accuracy. EFT was defined as the time at which the participant first fixated his or her gaze longer on the correct word choice than on the incorrect one. The difference between the proportion of time spent fixating on the correct answer and the incorrect answer, measured in 200ms bins, was calculated. EFT was considered the point in time when this proportion difference exceeded the threshold of 15% for 200ms or more (see Wendt et al., 2014). A graphical depiction of the average proportion of time spent fixating to the correct word is shown in Figure 1. ORT and EFT were measured from the sentence onset, allowing for comparisons between short and long separation sentences, as these sentences contained the same words in a different order and were therefore the same length. Since response times for incorrect responses are difficult to interpret, only ORTs and EFTs for correct trials were calculated and analyzed.

RESULTS

Accuracy

With a dependent variable of accuracy, the present study is a 2 x 3 within-subjects factorial design, with within-subjects factors of Syntactic Complexity (OR, SR) and Agency Distance (base, short, long). A within-subjects analyses of variance (ANOVA) was performed to compare these effects. Means and standard deviations for accuracy by sentence type can be found in Table 2. The proportion of correct comprehension responses for each type of sentence is depicted in Figure 2. Participants performed near ceiling for the syntactically simpler SR sentences in all three Agency Distance conditions. However, as expected, for sentences in the OR structure, while accuracy was still high, comprehension was not as successful. Specifically, there was poorer comprehension in OR long sentences compared to any other condition.

A significant main effect of Syntactic Complexity was found ($F[1, 117] = 19.105, p < .001, \eta_p^2 = .529$), replicating previous findings of poorer comprehension accuracy for the more demanding OR sentences than the simpler SR sentences. A significant main effect of Agency Distance was also found ($F[2, 34] = 13.753, p < .001, \eta_p^2 = .447$), such that long separation sentences were more often incorrect than base or short separation sentences. Additionally, there was a significant Complexity \times Agency Distance interaction ($F[2, 34] = 9.204, p = .001, \eta_p^2 = .351$), indicating that Agency Distance had a greater effect on the more complex OR sentences than the SR sentences. These findings reflect those of DeCaro et al. (2016).

A series of pairwise comparisons, corrected with Bonferroni, was performed to further analyze the differences in accuracy. Since SR accuracy was almost at ceiling for all levels of

agency distance, the following analyses were only conducted on OR conditions. Participants were incorrect significantly more on OR long sentences than OR base sentences ($p = .001$) and OR short sentences ($p = .005$). OR short and OR base sentences produced no significant differences in accuracy ($p = 1.000$). Accuracy was significantly worse in the OR short condition than the SR short condition ($p = .026$), as well as significantly worse in the OR long condition than the SR long condition ($p < .001$). Base sentences in each complexity resulted in similar accuracy ($p = .171$). These analyses indicate the interactive effect of complexity and agency distance, concluding that OR long sentences resulted in differentially poorer accuracy than any other type.

Response Time

Since response times were measured from the sentence onset, only short and long separation sentences were included in the following analyses because they were the same length (ten words), while base sentences only contained six words. Therefore, a $2 \times 2 \times 2$ within-subjects ANOVA was performed, with two levels of Response Type (EFT, ORT), two levels of Syntactic Complexity (SR, OR), and two levels of Agency Distance (short, long). Pairwise comparisons, corrected with Bonferroni, were also conducted to further examine these effects. ORTs and EFTs by sentence type are depicted graphically in Figure 3. Means and standard deviations for each sentence type can be found in Tables 3 and 4, respectively. For both response types, OR sentences overall took longer than SR sentences, but this gap narrowed for long separation sentences, which was surprising. We expected to find differentially longer EFTs and ORTs for OR long sentences.

There was a significant main effect of Response Type ($F[1, 17] = 322.816, p < .001, \eta_p^2 = .95$), indicating that, as expected, EFTs were significantly quicker than ORTs. There were also

significant main effects of Syntactic Complexity ($F[1, 17] = 47.328, p < .001, \eta_p^2 = .736$), such that OR sentences had slower responses than SR sentences, and Agency Distance ($F[1, 17] = 31.249, p < .001, \eta_p^2 = .648$), such that long separation sentences had slower responses than short separation sentences. Agency Distance had interactive effects with both Response Type ($F[1, 17] = 39.937, p < .001, \eta_p^2 = .701$) and Syntactic Complexity ($F[1, 17] = 20.965, p < .001, \eta_p^2 = .552$). Response Type and Syntactic Complexity did not interact ($F[1, 17] = 1.226, p = .284$). There was also a marginally significant three-way interaction ($F[1, 17] = 3.231, p = .09$).

From pairwise comparisons with the Bonferroni correction, in conjunction with Figure 3, it can be seen that the Agency Distance \times Syntactic Complexity interaction manifested in OR sentences taking longer than SR sentences in short separation trials ($p < .001$), and this difference disappearing in long separation trials ($p = .134$). It is surprising that SR long and OR long sentences produced similar response times, as it was expected that OR long sentences would take differentially longer than the other sentence types. Also, short separation sentences had significantly quicker response times than long separation sentences in the SR condition ($p < .001$), while there was no difference between them in the OR condition ($p = .684$). It is unclear why OR long sentences showed similar performance to OR short sentences. The Agency Distance \times Response Type interaction can be seen in ORT, where long separation sentences took significantly longer than short separation sentences ($p < .001$), while there was no significant difference between them in EFT ($p = .673$). This demonstrates that ORTs showed exaggerated effects of Agency Distance compared to EFTs, as hypothesized. Also, as expected with an on-line measure, EFTs were significantly quicker than ORTs in both short separation ($p < .001$) and long separation ($p < .001$) sentences.

Working Memory Span and Inhibitory Control Ability

A series of zero-order and partial correlations was conducted, examining how WM span, as indicated by participants' Reading Span score, and inhibitory control ability, as measured by conflict cost in the Flanker task, are related to sentence types and accuracy, EFTs, and ORTs. A higher Reading Span score indicated greater WM span, while a higher conflict cost value indicated lower inhibitory control ability. Zero order correlations and partial correlations for accuracy, ORTs, and EFTs can be found in Tables 5, 6, and 7, respectively. Interestingly, only relationships with EFTs reached significance at the .05 level, while some accuracy and ORT relationships with WM span were marginally significant.

Accuracy and WM span had a marginally significant correlation of .435 for OR long sentences ($p = .072$). When controlling for inhibitory control ability, this partial correlation increased slightly to .450 ($r = .070$). All other correlations involving accuracy were non-significant ($p > .100$). These findings suggest that it may be possible for successful sentence comprehension of complex sentences to occur even with impairments in WM span or inhibitory control ability, but this claim requires further examination.

ORT had a marginally significant negative correlation with WM span for SR long sentences ($r = -.426, p = .078$), suggesting a quicker response time for higher WM span. Controlling for inhibitory control ability, this negative partial correlation was marginally significant ($r = -.418, p = .095$). For OR long sentences, the negative partial correlation between ORT and WM span, controlling for inhibitory control ability, was marginally significant as well ($r = -.448, p = .071$), suggesting a similar pattern as in SR long sentences. No correlations between ORT and inhibitory control ability reached significance ($p > .100$). These findings suggest that the long separation sentences do in fact tax working memory resources, as greater

WM span was related to quicker ORTs in the long separation trials of both syntactic complexities, though the significance of these findings was marginal.

EFT significantly correlated with WM span for SR long sentences, ($r = -.474, p = .047$). This strong negative correlation suggests that a higher WM span was related to quicker EFTs in SR long sentences, and partially correlating these factors, controlling for inhibitory control ability, also resulted in significance ($r = -.486, p = .048$). For OR long sentences, the negative partial correlation between EFT and WM span, controlling for inhibitory control ability, reached significance ($r = -.518, p = .033$), as did the positive partial correlation for EFT and inhibitory control ability controlling for WM span ($r = .536, p = .027$). These correlations indicate that, in OR long sentences, a greater conflict cost, or a lower inhibitory control ability, was related to slower EFTs, while a greater WM span was related to quicker EFTs. Curiously, the partial correlation between EFT and WM span, controlling for inhibitory control ability, in SR short sentences neared significance ($r = -.438, p = .078$), suggesting a possible similar negative relation between WM span and EFT to SR long and OR long sentences, but this requires further investigation. It is not surprising that in both SR and OR long sentences, which tax working memory resources, higher WM span was related to quicker EFTs. Inhibitory control ability was only significantly related to EFT in OR long sentences, suggesting that reanalyzing the unexpected syntax of the OR sentence was particularly taxing in the long separation sentences.

DISCUSSION

This study is an early step in the process of understanding spoken language comprehension using eye-tracking. The methodology utilized in the present study is a novel paradigm, allowing for the on-line examination of whole-sentence comprehension using all three measures of accuracy, ORT, and EFT with a visual setup of two word options. A similar eye-tracking paradigm for whole sentences with pictorial representations has been used by Wendt et al. (2014). Ayasse, Lash, & Wingfield (2016) focused on the comprehension of a specific word in the context of a sentence using eye-tracking and pictorial options, while DeCaro et al (2016) only examined accuracy using a manual button press response. Using an on-line method in conjunction with overt responses provides a new understanding of whole-sentence comprehension through a continuous, time-locked measure. In combination with printed words, the methods used here are unique and provide new insight into the process of speech comprehension.

The present study found that participants were incorrect more often on OR sentences than SR sentences, and even more so for OR long sentences, as has been previously demonstrated. Additionally, it was found that OR sentences took longer to comprehend than SR sentences, and that long separation sentences took longer than short separation sentences, but interactions also played a role. OR short sentences resulted in quicker comprehension than SR short sentences, but OR long and SR long sentences surprisingly showed similar response times. Also, SR long sentences produced slower response times than SR short sentences, but OR short and OR long sentences were similar in response times. As it was expected that OR long sentences would take

differentially longer than other sentence types, it is surprising that OR long sentences had similar response times to both SR long and OR short. This could be due to the timing of the measure of response time in combination with the order of the words heard; in each sentence type, the two possible answers and the verb were heard at different points in the sentence. Since response time was measured from the sentence onset, it may have been affected by the varying points at which participants could likely figure out the answer (knowledge point) for each type of sentence. However, since the sentences differed in the order of the presented key words (i.e. noun-verb-noun vs. noun-noun-verb), measuring from the sentence onset was decided as the best option. Further research should examine response times measured from different points throughout the sentences.

EFTs did not show differences between long and short separation sentences, but ORTs showed slower response times for long separation sentences. These findings support the idea that EFT tells a different story than ORT. Referring back to Figure 3 and the significant interaction between Response Type and Agency Distance, it is clear that the effects of Agency Distance, and therefore working memory, seem to be exaggerated for ORT. This suggests that ORT relies on additional processes compared to EFT, and that EFT may in fact be a more accurate representation of the actual moment of comprehension than ORT. Furthermore, a distinction between accuracy and EFT was made, supporting the claim that EFT creates a different and potentially clearer look at the process of speech comprehension.

From the partial correlations, it is clear that WM span plays a major role in the speed of speech comprehension. While accuracy was generally unrelated to WM span and inhibitory control ability, ORT showed some weak relations to WM span. EFT had fairly strong partial correlations with WM span and inhibitory control ability for OR long sentences as well as with

WM span for SR long sentences. These results suggest that WM span, and possibly inhibitory control ability, are more closely related to the on-line interpretive processing of the sentence, as indicated by EFT, than the post-interpretive decision-making process, as measured by ORT. They also suggest that long separation sentences tax working memory resources in a way that is distinct from the base and short sentences, regardless of whether they were SR or OR. While manipulating syntactic complexity using SR and OR sentences can alter the difficulty of the sentences, the present results do not necessarily indicate that OR sentences themselves inherently tax working memory, as some literature has argued (Ferreira et al., 1996; Cooke et al., 2002; Just & Carpenter, 1992; Wingfield et al., 2006; Zurif et al., 1995). The findings suggest that WM span does not predict EFT differently for SR and OR sentences, and thus syntactic analysis is not affected by WM span. The effects of inhibition were less consistent and therefore require further research; it is possible that the Flanker task used, which only had 16 trials, is not sensitive enough to distinguish differences in inhibitory control ability.

The present results provide partial support for Caplan and Waters' model. As determining the agent of the action is considered by Caplan and Waters to be an interpretive process, and eye-fixations are an on-line measure that can capture this process, the lack of effect of WM span on EFT based on syntactic complexity supports their claim; even in more syntactically complex sentences, WM span did not predict EFT. However, the present study found a relationship between agency distance and WM span in EFT. WM span predicted EFT based on agency distance, suggesting that working memory resources were indeed required for the interpretive task of determining the agent.

Understanding how complex sentences and increased working memory demands affect younger adults' comprehension allows future researchers to make comparisons to older adults.

While older adults are graced with relatively preserved spoken language comprehension, (Wingfield, Tun, & McCoy, 2005; Helfer & Staub, 2014; Wingfield & Lash, 2016), they experience a number of declines, including poorer episodic memory, reduced working memory capacity, and slower processing speed as they age (Wingfield & Kahana, 2002; Salthouse, 1994, 2004; Salthouse & Babcock, 1991; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). Additionally taxing perceptual resources, as is done by the prevalent condition of age-related hearing loss, further taxes cognitive resources because these draw from the same pool of resources. Therefore, the effects found in younger adults may be larger and more exaggerated in older adults. However, considering the findings of Ayasse, Lash, and Wingfield (2016), these cognitive domains may not be inherently needed in order to comprehend speech successfully and may only influence post-interpretive processing. Studying the unique declines and cognitive changes among older adults using an on-line measure such as eye-tracking can contribute to the understanding of more complex speech comprehension in populations with limited cognitive abilities. Currently, the manipulations of Syntactic Complexity and Agency Distance are being tested on older adults with good and poor hearing, which will provide insight to the anticipated age effects and any additional effects of hearing acuity on comprehending spoken language.

There are some limitations to the present study. First, sentences were played in quiet and were individually adjusted to suprathreshold levels. In the real world, speech occurs at a wide range of volumes in the presence of near-constant background noise, so this experiment sets up an ideal environment for listening and understanding sentences. Additionally, as the sentence stimuli followed a limited number of forms, participants may have developed strategies for predicting the “doer” of the action as a result of repeated exposure to the sentence structure. Fillers of different syntactic structures were included in the stimuli to minimize the formation of

strategies, but participants were limited in the amount of time they could spend listening to sentences for practical reasons as well as risk of fatigue.

While the sentences used were not of the simplest possible form, speech in everyday context is much more varied. This study indicates comprehension of somewhat predictable, grammatically correct sentences in an ideal listening environment. Further studies should utilize this new paradigm for sentences in background noise, at varying presentation levels, and in additional syntactic structures. As impaired hearing and the presence of background noise can make understanding speech an especially challenging task for older adults (Wingfield, Tun, & McCoy, 2005; Helfer & Staub, 2014; Wingfield & Lash, 2016), the effects of these manipulations should be examined in participants of a wide age range.

CONCLUSION

Along with successfully testing a novel paradigm, the present study provides a new perspective on the process of speech comprehension in young adults. Upon examining accuracy, ORT, and EFT of sentences varying in Syntactic Complexity and Agency Distance, there is support that EFT is a more accurate indicator of the speed of sentence comprehension than off-line measures. The role of WM span in on-line response times suggests that complex sentences themselves do not tax working memory resources, supporting Caplan and Waters' model, but specifically manipulating working memory does in fact impact eye-gaze. It is not yet known how additional taxing of working memory resources, such as hearing loss, affects EFT. Further testing of this paradigm needs to be done, as well as varied syntactic and presentation manipulations, in order to better understand the process of comprehending speech throughout the lifespan.

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.
- Ayasse, N.D., Lash, A., & Wingfield, A. (2016). Effort not speed characterizes comprehension of spoken sentences by older adults with mild hearing impairment. *Frontiers in Aging Neuroscience*, *8*(329). doi: 10.3389/fnagi.2016.00329
- Baddeley, A.D. & Hitch, G.J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485-493.
- Caplan, D., Dede, G., Waters, G., Michaud, J., & Tripodis, Y. (2011.) Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging*, *26*(2), 439-450. doi: 10.1037/a0021837
- Caplan, D., Michaud, J., & Hufford, R. (2013). Short-term memory, working memory, and syntactic comprehension in aphasia. *Cognitive Neuropsychology* *30*(2), 77-109. doi: 10.1080/02643294.2013.803958
- Caplan, D. & Waters, G.S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 77-126.
- Caplan, D. & Waters, G.S. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin and Review*, *20*(2), 243-268. doi: 10.3758/s13423-012-0369-9
- Cooke, A., Zurif, E.B., De Vita, C., Alsop, D., Koenig, P., Detre, J., ... Grossman, M. (2002). Neural basis for sentence comprehension: Grammatical and short-term memory components. *Human Brain Mapping*, *15*, 80-94. doi: 10.1002/hbm.10006.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84-107.
- Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466.
- DeCaro, R., Peelle, J.E., Grossman, M., & Wingfield, A. (2016). The two sides of sensory-cognitive interactions: Effects of age, hearing acuity, and working memory span on sentence comprehension, *7*(236). doi: 10.3389/fpsyg.2016.00236
- Eriksen, B.A. & Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a nonresearch task. *Perception and Psychophysics*, *16*(1), 143-149.
- Ferreira, F., Henderson, J.M., Anes, M.D., Weeks, P.A., & McFarlane, D.K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 324-335. doi: 10.1037/0278-7393.22.
- Gibson, E., Bergen, L., & Piantadosi, S.T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences U.S.A.*, *110*, 8051-8056. doi: 10.1073/pnas.1216438110
- Hadar, B., Skrzypek, J.E., Wingfield, A., & Ben-David, B.M. (2016). Working memory load affects processing time in spoken word recognition: Evidence from eye-movements. *Frontiers in Neuroscience*, *10*(21). doi: 10.3389/fnins.2016.00221
- Helfer, K.S. & Staub, A. (2014). Competing speech perception in older and younger adults: Behavioral and eye movement evidence. *Ear Hear*, *35*(2), 161-170. doi: 10.1097/AUD.0b013e3182a830cf

- Hetherington, R. (1954). The Snellen Chart as a test of visual acuity. *Psychologische Forschung*, 24, 349-357.
- Holladay, J.T. (2004). Visual acuity measurements. *Journal of Cataract and Refractive Surgery*, 30, 287-290.
- Huetting, F. & McQueen, J.M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57, 460-482. doi: 10.1016/j.jml.2007.02.001
- Huetting, 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. doi: 10.1016/j.actpsy.2010.11.003
- Just, M.A. & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122-149.
- Katz, J. (2002). *Handbook of Clinical Audiology*, 5th Edn. Philadelphia, PA: Lippincott Williams & Wilkins.
- McCabe, D.P., Roediger III, H.L., McDaniel, M.A., Balota, D.A., & Hambrick, D.Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, 24(2), 222-243. doi: 10.1037/a0017619
- McQueen, J.M. & Viebahn, M.C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *The Quarterly Journal of Experimental Psychology*, 60(5), 661-671. doi: 10.1080/17470210601183890
- Novick, J.M., Trueswell, J.C., & Thompson-Schill, S.L. (2005). Cognitive control and parsing: Re-examining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 263-281. doi:10.3758/CABN.5.3.263
- Salthouse, T.A. (1994). Aging associations: Influence of speed on adult age differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1486-1503.
- Salthouse, T.A. (2004). What and when of cognitive aging. *Current Directions in Psychological Sciences*, 13(4), 140-144.
- Salthouse, T.A. & Babcock, R.L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, 27(5), 763-776.
- Salverda, A.P. & Tanenhaus, M.K. (2010). Tracking the time course of orthographic information in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5), 1108-1117. doi: 10.1037/a0019901
- Tanenhaus, M.K., Magnuson, J.S., Dahan, D. & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, 29(6), 557-580.
- Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *American Advancement for the Advancement of Science*, 268(5217), 1632-1634.
- Tun, P.A., Benichov, J., & Wingfield, A. (2010). Response latencies in auditory sentence comprehension: Effects of linguistics versus perceptual challenge. *Psychology and Aging*, 25(3), 730-735. doi: 10.1037/a0019300

- Waters, G.S. & Caplan D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging, 16*(1), 128-144. doi: 10.1037/0882-7974.16.1.128
- Wendt, D., Brand, T., & Kollmeier, B. (2014). An eye-tracking paradigm for analyzing the processing time of sentences with different linguistic complexities. *PLoS ONE, 9*(6). doi: 10.1371/journal.pone.0100186
- Wendt, D., Dau, T., & Hjortkjaer, J. (2016). Impact of background noise and sentence complexity on processing demands during sentence comprehension. *Frontiers in Psychology, 7*(345). doi: 10.3389/fpsyg.2016.00345
- Wendt, D., Kollmeier, B., & Brand, T. (2015). How hearing impairment affects sentence comprehension: Using eye fixations to investigate the duration of speech processing. *Trends in Hearing, 19*, 1-18. doi: 10.1177/23312165155584149
- Wingfield, A. & Grossman, M. (2006). Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging. *Journal of Neuropsychology, 96*(6), 2830-2839. doi: 10.1152/jn.00628.2006
- Wingfield, A. & Kahana, M.J. (2002). The dynamics of memory retrieval in older adulthood. *Canadian Journal of Experimental Psychology, 56*(3), 187-199. doi: 10.1037/h0087396
- Wingfield, A., & Lash, A. (2016). Audition and the comprehension of spoken language in adult aging: Stability in the face of change. In K.W. Schaie & S. Willis (Eds.), *Handbook of the Psychology of Aging* (8th ed.; pp. 165–185). London: Elsevier.
- Wingfield, A., McCoy, S.L., Peelle, J.E., Tun, P.A., & Cox, L.C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *Journal of the American Academy of Audiology, 17*, 487-497. doi: 10.3766/jaaa.17.7.4.
- Wingfield, A., Peelle, J.E. & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging Neuropsychology and Cognition, 10*(4), 310-322.
- Wingfield, A., Tun, P.A., & McCoy, S.L. (2005). Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Current Directions in Psychological Science, 14*(3), 144-148.
- Yoon, J., Campanelli, L., Goral, M., Marton, K., Eichorn, N., & Obler, L.K. (2015). The effect of plausibility on sentence comprehension among older adults and its relation to cognitive functions. *Experimental Aging Research, 41*(3), 272-302, doi: 10.1080/0361073X.2015.1021646
- Zachary, R.A. (1986). *Shipley Institute of Living Scale: Revised Manual*. Los Angeles, CA: Western Psychological Services.
- Zurif, E., Swinney, D., Prather, P., Wingfield, A., & Brownell, H. (1995). The allocation of memory resources during sentence comprehension: evidence from the elderly. *Journal of Psycholinguistic Research, 24*, 165-182. doi: 10.1007/BF02145354.

Table 1.

Sample Sentences of Each Type

Type	Example
SR base	Girls that confuse boys are mean.
SR short	Girls that confuse boys in blue corduroy vests are mean.
SR long	Girls in blue corduroy vests that confuse boys are mean.
OR base	Boys that girls confuse are mean.
OR short	Boys in blue corduroy vests that girls confuse are mean.
OR long	Boys that girls in blue corduroy vests confuse are mean.

Table 2.

Means and Standard Deviations for Accuracy by Sentence Type

Sentence Type	Mean (proportion correct)	SD
SR base	.987	.029
SR short	.988	.028
SR long	.991	.025
OR base	.959	.075
OR short	.945	.069
OR long	.844	.133

Table 3.

Means and Standard Deviations for Overt Response Times by Sentence Type in Seconds

Sentence Type	Mean (s)	SD
SR base	2.467	.516
SR short	2.880	.884
SR long	3.785	.479
OR base	2.874	.645
OR short	3.680	.621
OR long	3.908	.896

Table 4.

Means and Standard Deviations for Eye-Fixation Times by Sentence Type in Seconds

Sentence Type	Mean (s)	SD
SR base	.795	.356
SR short	.858	.409
SR long	1.066	.581
OR base	1.015	.376
OR short	1.426	.604
OR long	1.265	.657

Table 5.

Zero-Order and Partial Correlations for (a) Accuracy, (b) Working Memory Span, and (c) Inhibitory Control Ability

	SR base	SR short	SR long	OR base	OR short	OR long
Zero-order						
Correlations (<i>r</i>)						
ab	.140	-.024	-.134	.180	.197	.435 [^]
ac	.269	-.364	-.287	-.178	.094	.014
Partial						
Correlations (<i>pr</i>)						
ab.c	.067	.092	-.055	.247	.178	.450 [^]
ac.b	.240	-.373	-.262	-.246	.039	-.132
<i>Note.</i> [^] $p < .1$, * $p < .05$, ** $p < .01$						

Table 6.

Zero-Order and Partial Correlations for (a) Overt Response Time, (b) Working Memory Span, and (c) Inhibitory Control Ability

	SR short	SR long	OR short	OR long
Zero-order				
Correlations (<i>r</i>)				
ab	-.355	-.426 [^]	-.317	-.371
ac	-.166	-.098	-.033	.173
Partial				
Correlations (<i>pr</i>)				
ab.c	-.325	-.418 [^]	-.322	-.448 [^]
ac.b	-.070	-.031	.066	.317
<i>Note.</i> [^] $p < .1$, * $p < .05$, ** $p < .01$				

Table 7.

Zero-Order and Partial Correlations for (a) Eye-Fixation Time, (b) Working Memory Span, and (c) Inhibitory Control Ability

	SR short	SR long	OR short	OR long
Zero-order				
Correlations (<i>r</i>)				
ab	-.349	-.474*	-.248	-.347
ac	.210	-.035	.210	.379
Partial				
Correlations (<i>pr</i>)				
ab.c	-.438 [^]	-.486*	-.331	-.518*
ac.b	.348	.123	.305	.536*

Note. [^] $p < .1$, * $p < .05$, ** $p < .01$

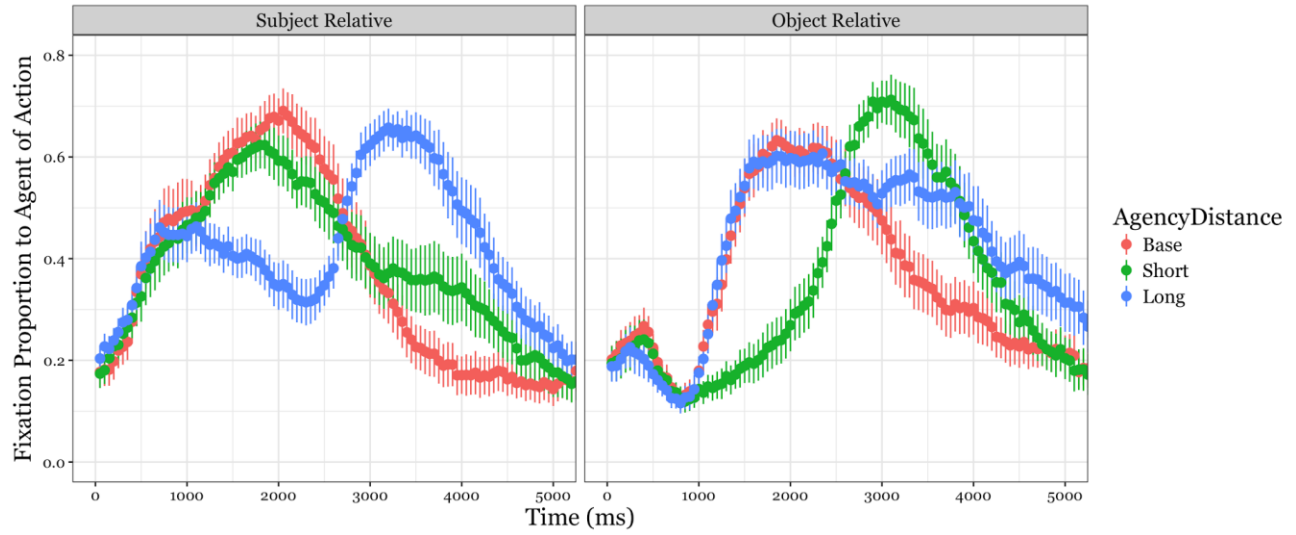


Figure 1. Average proportion of time spent fixating to the agent or “doer” in each sentence type. Time 0 indicates the sentence onset. Error bars represent standard errors of the mean.

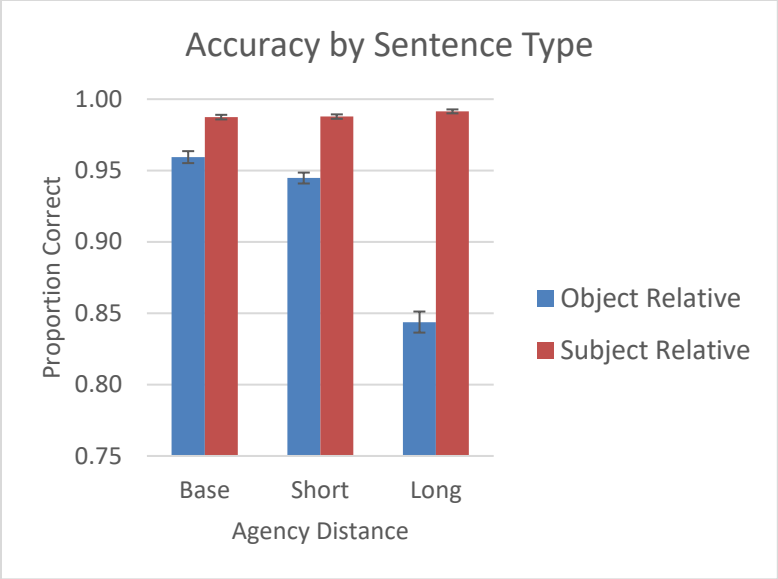


Figure 2. Accuracy by complexity and agency distance. Error bars represent standard errors of the mean.

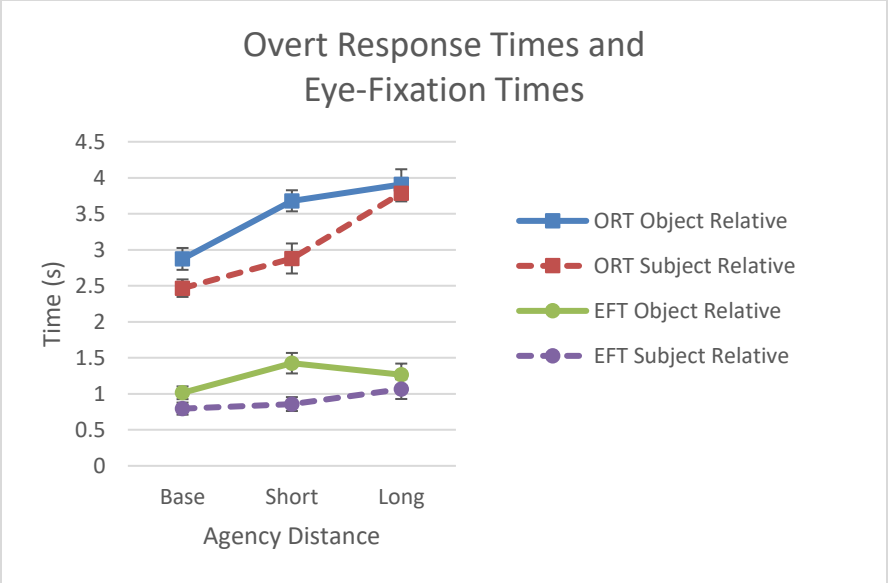


Figure 3. Overt response time and eye-fixation time by complexity and agency distance. Error bars represent standard errors of the mean.