Do resource constraints affect lexical processing?

Evidence from eye movements

Mallorie Leinenger[a], Mark Myslín[b], Keith Rayner[a], & Roger Levy[b]

[a]Department of Psychology, [b]Department of Linguistics

University of California, San Diego

Correspondence to:

Mallorie Leinenger

Department of Psychology

Denison University

100 West College Street

Granville, Ohio 43023

Email: leinengerm@denison.edu

Abstract

Human language is massively ambiguous, yet we are generally able to identify the intended meanings of the sentences we hear and read quickly and accurately. How we manage and resolve ambiguity incrementally during real-time language comprehension given our cognitive resources and constraints is a major question in human cognition. Previous research investigating resource constraints on lexical ambiguity resolution has yielded conflicting results. Here we present results from two experiments in which we recorded eye movements to test for evidence of resource constraints during lexical ambiguity resolution. We embedded moderately biased homographs in sentences with neutral prior context and either long or short regions of text before disambiguation to the dominant or subordinate interpretation. The length of intervening material had no effect on ease of disambiguation. Instead, we found only a main effect of meaning at disambiguation, such that disambiguating to the subordinate meaning of the homograph was more difficult—results consistent with the reordered access model and contemporary probabilistic models, but inconsistent with the capacity-constrained model.

Keywords: lexical ambiguity, digging-in effects, eye movements, reading

During normal reading, readers make use of contextual information to help them resolve ambiguities inherent in the text. However, not all information relevant for disambiguation is available immediately, such that, in some cases readers encounter ambiguities for which the intended meaning is unknown given the lack of contextual cues. One such type of ambiguity is LEXICAL AMBIGUITY – for example, one sense of the word *wire* could be paraphrased "thin metal filament", another as "telegram". Although leading models of lexical ambiguity resolution agree that readers are able to make use of available contextual information to help them activate and integrate the appropriate meaning of an ambiguous word, they disagree on exactly what readers do in situations where the intended meaning is unknown. In the absence of contextual disambiguating information, readers must either maintain multiple meanings of an ambiguous word, or select one meaning to elaborate and integrate. In the latter case, if they select the meaning ultimately intended in the discourse, reading can continue without disruption, but if they have chosen to integrate the incorrect meaning, disruption and reanalysis may occur when later context disambiguates toward the unselected meaning. Thus, it may be advantageous for a reader to maintain multiple meanings in parallel if they are able. However, maintaining multiple meanings may be substantially SUBJECT TO RESOURCE CONSTRAINTS: maintaining multiple meanings might require or tax scarce cognitive resources such as limited working memory, and prolonged maintenance may well be impossible.

To the extent that readers do attempt to maintain multiple meanings of ambiguous words encountered in neutral contexts and maintenance of multiple meanings is subject to resource constraints, we might expect to find that readers are unable to maintain all meanings at a steady level of activation over time. The activation of less preferred meanings might decrease, making resolution to less preferred meanings more effortful the longer the ambiguity persists before

disambiguation—so called DIGGING-IN EFFECTS (e.g., Levy, Reali, & Griffiths, 2009; Tabor & Hutchins, 2004). Such digging-in effects have primarily been documented and discussed in the context of syntactic ambiguity resolution (e.g., during the reading of garden path sentences— Ferreira & Henderson, 1991, 1993; Frazier & Rayner, 1982; Tabor & Hutchins, 2004; see also General Discussion). It is less clear, however, whether lexical ambiguity resolution is subject to such resource constraints. Here we report two experiments bearing on this question.

Depending on the nature of and constraints on the processing of lexical ambiguities, three distinct possibilities emerge for the resolution of ambiguous words encountered in neutral contexts: (1) readers do not attempt to maintain multiple meanings of an ambiguous word, (2) readers maintain multiple meanings of an ambiguous word and such maintenance is not subject to resource constraints, or (3) readers maintain multiple meanings of an ambiguous word, but such maintenance is subject to resource constraints. The third possibility predicts that digging-in effects should be observed at disambiguation (because the strength of a given meaning's activation will decay as cognitive resources are depleted), whereas the first two possibilities predict that digging-in effects should not be observed (either because only one meaning is rapidly selected or because multiple meanings are maintained across time without significantly taxing cognitive resources).

The question of whether or not digging-in effects are observed during lexical ambiguity resolution is important for distinguishing between leading models of lexical ambiguity resolution, which make different predictions regarding the presence (or absence) of such effects. Therefore we next review relevant data and theory in lexical ambiguity resolution, contrasting two classes of models: exhaustive access models that do not predict that digging-in effects should be observed (such as the reordered access model—Duffy, Morris, & Rayner, 1988) and

memory-based models that do (such as the capacity-constrained model—Miyake, Just, & Carpenter, 1994).

EXHAUSTIVE ACCESS MODELS assume that all meanings of a word are initially activated. In many exhaustive access models, it is additionally assumed that one meaning is rapidly integrated into the sentence. Indeed, there is evidence from offline tasks (e.g., cross-modal priming) that immediately after encountering an ambiguous word, multiple meanings may be active, but one meaning is selected relatively quickly (typically on the order of 100-200ms) with other meanings decaying or being actively inhibited (e.g., Onifer & Swinney, 1981; Simpson, 1981; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979). Even without strongly biasing context preceding an ambiguous word, evidence from cross-modal priming suggests that a weakly dispreferred meaning may not be maintained for more than a few syllables after the word is encountered (e.g., Swinney, 1979; but see Hudson & Tanenhaus, 1984 for evidence that multiple meanings might be maintained for slightly longer in the absence of prior disambiguating context). Further evidence for exhaustive access comes from online reading, where, in the absence of prior disambiguating context, readers look longer at balanced homographs (words with two equally frequent meanings) than at unambiguous control words matched for length and frequency (e.g., Duffy et al., 1988; Rayner & Duffy, 1986).

The REORDERED ACCESS MODEL (Duffy et al., 1988), a more specific exhaustive access model, further specifies the interaction of sentence context and relative meaning frequency. According to this model, upon first encountering an ambiguous word there is immediate access to the alternative meanings based on their frequency and the bias of the preceding context, competition among the meanings ensues, and resolving that competition takes longer the more equal the initial footing among the alternative meanings. Hence, in a neutral prior context a

balanced homograph will be read more slowly than a matched control word, but a biased homograph (a word with two meanings of differing frequency, a higher frequency DOMINANT meaning and a lower frequency SUBORDINATE meaning) is read more or less as quickly as a matched control word (Simpson & Burgess, 1985), since the homograph's dominant meaning is activated first and easily integrated into the sentence. Following a strongly biasing prior context, a biased homograph will be read quickly when its dominant meaning is the one more consistent with the context, but will be read more slowly when its subordinate meaning is the one more consistent with the context, as competition among the two meanings is prolonged (Duffy et al., 1988; Pacht & Rayner, 1993; Sheridan & Reingold, 2012; Sheridan, Reingold, & Daneman, 2009—but see e.g., Colbert-Getz & Cook, 2013; Leinenger & Rayner, 2013; Wiley & Rayner, 2000, for evidence that this so-called SUBORDINATE BIAS EFFECT can be reduced). In sum, in the absence of contextual information, the two meanings of a homograph are accessed in the order of their meaning frequency, with the dominant meaning being accessed first and integrated into the sentence very rapidly (e.g., Rayner, Cook, Juhasz, & Frazier, 2006). In the reordered access model, readers do not attempt to maintain multiple meanings even following neutral contexts, so lexical ambiguity does not increase cognitive resource requirements (consistent with theoretical possibility (1) above). If a homograph is preceded by neutral context and later disambiguated, this model predicts that disambiguation to the subordinate meaning will be more difficult than disambiguation to the dominant meaning regardless of how much material intervenes between the homograph and disambiguation.

Indeed, Rayner and Frazier (1989, Experiment 1) reported this pattern of results for highly biased homographs preceded by neutral context and disambiguated either immediately or three to four words later, further suggesting that for highly biased homographs, meaning

selection is immediate and multiple meanings are not maintained. Examples of the short and long subordinate-resolution conditions from Rayner and Frazier (1989, Experiment 1) are shown in (1) and (2), with short and long dominant-resolution conditions in (3) and (4). In these examples, the homograph is underlined and the disambiguating word is italicized.

    (1) George said that the <u>wire</u> *informed* John that his aunt would arrive on Monday.

    (2) George said that the <u>wire</u> was supposed to *inform* John that his aunt would arrive on Monday.

    (3) George said that the <u>wire</u> *surrounded* the entire barracks including the rifle range.

    (4) George said that the <u>wire</u> was supposed to *surround* the entire barracks including the rifle range.

Rayner and Frazier reported that GAZE DURATIONS (the sum of all initial fixation durations on a region before leaving it), measured in milliseconds per character, were consistently longer at the disambiguating word in resolution-to-subordinate-meaning conditions (1)—(2) than in resolution-to-dominant-meaning conditions (3)—(4), and did not change with presence or absence of an intermediate neutral region. Although this is initial suggestive evidence against digging-in effects in lexical ambiguity resolution, its interpretation is limited by the fact that the disambiguating word has a different form in each of the four conditions.

    A different class of exhaustive access models, PROBABILISTIC RANKED-PARALLEL MODELS, also predicts the absence of digging-in effects, not because only one meaning is maintained, but because under such models, readers are able to maintain multiple meanings without significantly taxing scarce cognitive resources (consistent with theoretical possibility (2)). For example, in syntactic parsing, many models propose probabilistic, parallel

disambiguation, such as the SURPRISAL model of Hale (2001) and Levy (2008). In this model, multiple syntactic parses are maintained, ranked according to the likelihood of each parse given the preceding sentence context, which is updated as new information is read. Such parallel, probabilistic models can easily be extended to lexical ambiguity resolution (e.g., Jurafsky, 1996, 2003). The simplest instantiations of these models allows unlimited maintenance in parallel, without cost, of all interpretations that have not been ruled out by incompatible input. These simplest instantiations thus predict that, following neutral context, disambiguation will be equally effortful regardless of how much neutral material intervenes between the homograph and disambiguation. When an ambiguous word is encountered in a neutral context, each meaning will be maintained with strength (formally, probability) roughly that of each meaning's frequency. Upon encountering disambiguating material, the different probabilities for the ambiguous word's meaning will affect processing difficulty: all else being equal, the disambiguating material's surprisal (log-inverse probability) will generally be higher the lower the probability of the compatible meaning. Thus disambiguation to a subordinate meaning will be higher surprisal, and therefore more difficult, than disambiguation to a dominant meaning, but the simplest instantiations of these models do not predict digging-in effects since multiple meaning maintenance is not subject to appreciable resource constraints.

Alternatively, the memory-oriented CAPACITY-CONSTRAINED MODEL of Miyake et al. (1994) assumes that readers might attempt to maintain multiple meanings of an ambiguous word, and that their working memory capacity for language constrains the degree to which they are able to do so (consistent with theoretical possibility (3)). According to the capacity-constrained model, working memory is a general "computational arena" that subserves both lexical processing and storage of intermediate and final products of comprehension. Meanings of

ambiguous words are activated in parallel, with more frequent meanings being activated more quickly and to a greater extent. Following a strong biasing context, only the supported meaning is integrated into the sentence and the other meaning either decays or is actively suppressed. In the absence of preceding disambiguating information (i.e., following a neutral context), readers may create dual mental representations, elaborating them in parallel until disambiguating information is reached. This maintenance and elaboration is subject to resource constraints, such that activation of the different meanings is reduced as memory capacity fills (e.g., with increasing words being read, or for readers with smaller working memory capacities), such that with time, the activation of subordinate meanings may fall below threshold. This model thus predicts an interaction of meaning dominance and resource availability: subordinate meanings should be especially difficult to process when few resources are available. Furthermore, since initial meaning activation varies as a function of meaning frequency, the activation of the subordinate meanings will persist longest (relative to the dominant meaning) when the homograph in question is balanced, and will fall below threshold faster with decreasing subordinate meaning frequency (i.e., for highly biased homographs).

Indeed, Miyake et al. (1994) supported this prediction for moderately biased homographs in two self-paced reading experiments with separate manipulations of memory and resource availability. First, they showed that readers with low working memory span are especially slow to process subordinate resolutions of ambiguous words (Miyake et al., 1994, Experiment 1). Second, they manipulated the length of material intervening between subordinate homographs and eventual disambiguation, creating short and long subordinate conditions, examples of which are shown in (5) and (6), with the homograph underlined and the disambiguating region italicized (Miyake et al., 1994, Experiment 2).

(5) Since Ken liked the <u>boxer</u>, he went to the *pet store to buy the animal.*

(6) Since Ken liked the <u>boxer</u> very much, he went to the nearest *pet store to buy the*

animal.

With this length manipulation they showed that mid-span readers exhibit digging-in effects in lexical ambiguity resolution, processing subordinate resolutions especially slowly when additional words intervene before disambiguation. Thus, their results run counter to those of Rayner and Frazier (1989) who found only an effect of meaning at disambiguation.

There are a few notable differences between the stimuli and methods used by Rayner and Frazier (1989) and Miyake et al. (1994) First, the homographs used by Rayner and Frazier were highly biased, such that the average probability of picking the dominant meaning in offline norming was .92 (range .76-1). In contrast, Miyake et al. used moderately biased homographs where the dominant to subordinate frequency ratio was judged to be 7.8:2.2. As already stated, according to the capacity-constrained model, the subordinate meaning falls below threshold faster with increasing frequency disparity. Thus, the failure of Rayner and Frazier to find an interaction of meaning frequency and length to disambiguation may have been due to the nature of their targets. In other words, it may be the case that the subordinate meaning frequency for their targets was so low that the subordinate meaning fell below threshold as early as the next word in the sentence (i.e., the point of early disambiguation). In contrast, for Miyake et al.'s moderately biased homographs, the subordinate meaning may have persisted until disambiguation in their shorter sentences (and for readers with high working memory capacities), but fell below threshold before the point of disambiguation in the longer sentence versions (and for readers with low working memory capacity). Second, in Miyake et al.'s Experiment 2, where

memory capacity was held constant, critical comparisons were between the reading of a homograph (eventually disambiguated to its subordinate sense) and the same sentence frame with the homograph replaced with an unambiguous semantic associate (e.g., *wrestler* or *collie* in the case of *boxer*). This allowed for easy comparison of the disambiguating region, which was identical across conditions, but made comparisons of the critical word difficult, as they were not matched on lexical variables known to influence reading time. The words differed in length, and the authors did not specify whether these controls were frequency-matched, and if so, whether to the frequency of the homograph overall, the frequency of its dominant sense, or the frequency of its subordinate sense, three scenarios that can produce markedly different results (e.g., Sereno, O'Donnell, & Rayner, 2006). In contrast, Rayner and Frazier compared reading on the same critical homograph later disambiguated to either its dominant or subordinate sense. This avoided the issue of deciding which frequency to match the control word to and allowed for easy comparison of reading times on the homograph in different disambiguation conditions, but made comparisons of the disambiguating region slightly more challenging as it contained different words across conditions. Third, Rayner and Frazier's results were obtained using eye tracking, which allowed for millisecond precision in sampling the location of the eyes during reading— providing a very sensitive measure of online processing. In contrast, Miyake et al.'s use of self-paced reading obscured whether their reported length-based digging-in effect arose instantaneously upon reaching disambiguation, or was instead associated with clause wrap-up, since effects were observed primarily in sentence-final words that were within the spillover region following the disambiguating word. Finally, the lexical digging-in effect reported by Miyake et al. was significant by subjects only, raising questions about the reliability of the effect. Thus it remains an open question whether digging-effects manifest in lexical ambiguity

resolution within a sentence.

Interestingly, despite their differing architectures and assumptions, these models make similar behavioral predictions regarding how lexical ambiguity resolution proceeds in other situations. For example, both the reordered access model and the capacity constrained model agree that, following strong biasing context, only the contextually appropriate meaning of a homophone is integrated and maintained. These models only differ in their predictions surrounding behavior in the absence of prior disambiguating context, making it the critical test case to adjudicate between them.

In the current study, we sought to further test the predictions of the reordered access model[1] and the capacity-constrained model while also attempting to reconcile the different results obtained by Rayner and Frazier (1989) and Miyake et al. (1994). Consistent with Rayner and Frazier, we collected eye movement data to allow for a fine-grained investigation of potential digging-in effects in lexical ambiguity resolution. We used moderately biased homographs, consistent with Miyake et al., to determine whether digging-in effects might only be observed for less highly-biased homographs, for which the capacity-constrained model predicts both meanings can be maintained for some amount of time. We constructed our stimuli similarly to Rayner and Frazier, such that critical comparisons were between conditions where the homograph was ultimately disambiguated to either its dominant or subordinate meaning after a short or long intervening region of text (but see Experiment 2 for stimuli similar to those used by Miyake et al., 1994 in their Experiment 2).

The reordered access model and the capacity-constrained model agree that, at the point of

---

[1] Because the reordered access model and the probabilistic ranked-parallel models make similar behavioral predictions (i.e., that digging-in effects should not be observed), we will discuss the current study's implications for the reordered access model, but will return to a discussion of all models in the general discussion.

disambiguation, readers should not have trouble disambiguating a moderately biased homograph to its dominant meaning. They differ in their predictions regarding disambiguation to the subordinate meaning. The reordered access model predicts immediate meaning selection for the homograph even following a neutral prior context, such that the dominant meaning will be selected and integrated, making disambiguation to the subordinate meaning more difficult than disambiguation to the dominant meaning regardless of how much text intervenes between the homograph and subsequent disambiguation (measured as longer reading times on the disambiguating region and/or more regressions to reread the homograph or preceding text). In contrast, the capacity-constrained model predicts that both meanings will be maintained until resource limitations cause the subordinate meaning to fall below threshold, such that disambiguation to either meaning will be easy with short regions of intervening text between the homograph and disambiguation (though disambiguation to the subordinate meaning will still be more difficult than disambiguating to the more highly active dominant meaning), but disambiguation to the subordinate meaning becomes increasingly more difficult with increasing length of intervening text.

Experiment 1

Methods

     Participants. Sixty native English speakers from the University of California, San Diego received course credit for their participation in the study. All participants had normal or corrected-to-normal vision.

     Apparatus. Participants' eye movements were monitored using an Eyelink 1000 eyetracker, which sampled and recorded eye position every millisecond. Subjects were seated 61

cm away from a 19-inch ViewSonic LCD monitor. Text was displayed in 14- point, fixed-width Consolas font, and 4 characters equaled 1º of visual angle. Viewing was binocular with eye location sampled from the right eye.

Materials. Prior to material creation, thirty-three native English speakers from the United States participated in online norming through Amazon's Mechanical Turk service for monetary compensation. They were given a list of words, one at a time, and asked to construct sentences containing each word (one sentence per word). Prior to the start of the norming, each participant was shown two examples where sentences were constructed using each word in its noun sense. In this way we hoped to covertly bias the participants to compose sentences using the noun senses of our homographs (many of our homographs, such as *nail*, *buzz*, *pit*, *toast*, and *finish*, also have verbal uses, e.g., "Alex couldn't *toast* the bread because the power was out.", but we used exclusively noun senses in our materials) without expressly instructing them to do so and potentially highlighting the ambiguous nature of our stimuli. 80 homographs and 64 unambiguous words were included for a total of 144 words, and it took participants approximately forty minutes to compose sentences for all of the words. Sentences were then coded for which meaning was expressed, and the overall bias of each homograph was computed as the proportion of participants expressing one versus the other meaning for the homograph. Based on the results of this norming, we selected thirty-two ambiguous words for which the probability of generating the dominant meaning ranged from 0.56 to 0.87 (mean = 0.72) such that the homographs were moderately biased. We compared the bias ratings that we obtained via Amazon's Mechanical Turk with previous norms collected at the University of Alberta; only 26 of the homographs used in the current study are contained in the Alberta norms, but for those 26 homographs, our norms and the Alberta norms are highly correlated with each other ($r(24) = .47$,

p = .015; Twilley, Dixon, Taylor, & Clark, 1994).

Four sentence versions (dominant short, dominant long, subordinate short, subordinate long) were created for each of the thirty-two biased homographs, resulting in a total of 128 experimental sentences. Both long and short sentence versions contained regions of text between the homograph and disambiguation. We refer to this region of text in the short conditions as the INTERMEDIATE region; this region averaged 4.16 words in length. In the long conditions, this region of text consisted of the intermediate region plus a LENGTHENING region (average length 4.03 words) that was inserted either before the intermediate region or in the middle of the intermediate region, to increase the amount of material that subjects read prior to disambiguation. Material appearing before the lengthening region in the long conditions we refer to as Intermediate 1; material appearing after the lengthening region we refer to as Intermediate 2. Since one region, Intermediate 1, is not present in all items, we additionally defined an Intermediate Super-region comprising Intermediate 1 (if present), Lengthener (if present), and Intermediate 2. The Intermediate Super-region thus reflects, across all items, the totality of material intervening between the homograph and disambiguation. We defined the disambiguating region as extending from the first word that differed between the dominant and subordinate versions of a given sentence pair, to the end of the sentence.

We also identified a disambiguating word post-hoc, to facilitate comparisons with previous research. In order to do so, we gave an additional set of six Mechanical Turk subjects, who did not participate in the other set of online norming, the short versions of our stimuli and asked them to select the first word in each sentence that they believed disambiguated the homograph. This norming revealed that, for the majority of our stimuli (70.3%), subjects did not unanimously agree on which word was the first to disambiguate the homograph. We defined the

disambiguating word as the most commonly selected word in our norming[2]. Across all items, 77.3% of subject responses were in agreement with the word analyzed as the disambiguating word. It should be noted that these words were not intentionally matched in the design of the experiment; however, across all items, there were no significant differences between the dominant and subordinate disambiguating words in length or log-transformed HAL frequency[3] (both $ps > .17$). Sample stimuli appear in Table 1, and the full set of stimuli is listed in Appendix A.

Four experimental lists were constructed and counterbalanced such that within each list each condition appeared an equal number of times, and across experimental lists each sentence appeared once in each of the four conditions. The thirty-two experimental sentences were presented along with seventy filler sentences. Simple comprehension questions appeared after 16 of the critical items and 34 of the filler items. Meaning condition (dominant, subordinate) and length condition (short, long) were tested within participants; however, each participant saw only one sentence for each homograph. The beginning of the sentence (prior to the homograph) and the lengthening region (which followed the homograph in the long sentence versions) were always neutral, and the disambiguating region always supported only one meaning of the homograph.

Table 1

Sample Stimuli for Experiment 1

Meaning	Length	Intro	Homograph	Intermediate 1*	Lengthener *	Intermediate 2*	Disambiguation

[2] For 2 items (dominant sentence frames for "deed" and "yard"), there was a tie for the most commonly selected word in the online norming, so the first author chose from among the most common responses.

[3] HAL frequency information was unavailable for the exact forms of 7 words (2 dominant and 5 subordinate resolutions), including possessives such as *animal's* and compounds such as *B-flat*; these cases were excluded from this analysis.

Dominant	Short	The	toast	I made		was really	delicious with blackberry jam.
Dominant	Long	The	toast	I made	a few days ago	was really	delicious with blackberry jam.
Subordinate	Short	The	toast	I made		was really	eloquent and well-delivered.
Subordinate	Long	The	toast	I made	a few days ago	was really	eloquent and well-delivered.

Note. Asterisked regions are part of the Intermediate Super-region

Procedure. Each participant was run individually in a session that lasted approximately thirty minutes. At the start of the experiment, participants completed a calibration procedure by looking at a random sequence of three fixation points presented horizontally across the middle of the computer screen. Each trial required the participant to fixate a point in the center of the screen before moving their eyes to a black square (40 pixels wide and 40 pixels tall), which appeared on the left side of the screen after the central fixation mark disappeared. This box coincided with the left side of the first character of the sentence and once a stable fixation was detected within the box, the sentence replaced it on the screen.

Prior to the experimental portion, 10 practice sentences (with 6 comprehension questions) were presented. All sentences were randomized for each participant and vertically centered on the screen. Participants were instructed to read silently at their normal pace for comprehension, and to press a button on a keypad when they finished reading. When comprehension questions appeared on the screen after a sentence, participants were required to respond *yes* or *no* via button press. Following incorrect answers, the word *incorrect* was displayed for 3 seconds before the next trial was initiated. Following correct answers, there was no feedback and the experiment continued with the next trial. Participants were correct on an average of 95.9% of questions.

Eye movement measures. Both early and late eye movement measures for the target homograph, the neutral pre-disambiguating region, and the disambiguating region were assessed (Rayner, 1998, 2009). We present separate analyses of pre-disambiguation measures (all measures prior to fixating any portion of the disambiguating region) and post- disambiguation

measures (all measures after fixating any portion of the disambiguating region). Both the reordered access model and the capacity-constrained model predict that initial reading of the homograph, and any re-reading prior to disambiguation, should not differ as a function of meaning, since no disambiguating information has yet been encountered. Therefore, for pre-disambiguation we report FIRST PASS TIME (the sum of all initial fixation durations on a region before leaving it—for single-word regions, this is equivalent to gaze duration), SECOND PASS TIME (the sum of all fixations on a region, excluding first pass fixations, before any disambiguating material was fixated; a value of zero is recorded when no second pass time occurred), and REREADING TIME (calculated the same as second pass time, except that values of zero were excluded, making this a pure measure of rereading time deconfounded from the probability of rereading) on the homograph, first and second pass time on the lengthening region, and first pass time on the second intermediate region.[4] We also report the probability of making a first-pass regression out of the intermediate super-region (REGRESSION PROBABILITY), as well as the probability that a first-pass regression out of this intermediate region leads to a fixation (not necessarily in a single saccade) on the homograph before the first fixation rightward of the intermediate region or the end of reading (in general, we call these "X←Y REGRESSIONS"). Following disambiguation, the reordered access model predicts more difficulty processing subordinate resolutions regardless of length, whereas the capacity-constrained model predicts more difficulty processing subordinate resolutions, which increases with the length of intervening material. Processing difficulty could manifest as either longer first pass reading of the disambiguating region or more regressions back to and longer re-reading of the homograph or beginning of the sentence. Post-disambiguation, we thus report first pass time on the

---

[4] We could not assess second-pass time pre-disambiguation for the second intermediate region, since it was followed immediately by the disambiguating region.

disambiguating region (measured in ms/char since the text differed across meaning conditions), go-past time on the disambiguating region, second pass time (the sum of all fixations on a region, excluding first pass fixations, after having fixated disambiguating information), rereading time, and total time on the homograph, the probability of making a regression out of the disambiguating region, and the probability of an X←Y regression from the disambiguating region to the homograph. We also report three post-hoc analyses of first pass time, go-past time, and total time on the disambiguating word. Digging-in effects could manifest as interactions in any of these post-disambiguation measures between meaning and length conditions. Furthermore, main effects of meaning condition on second-pass time on the homograph, probability of regressions out of the disambiguating region, and/or probability of an X←Y regression from the disambiguating region to the homograph that emerged only after disambiguation, would suggest that disambiguating to one meaning was more difficult than the other.

Results

Prior to analysis, fixations under 81ms were deleted, or pooled if they were within 1 character of another fixation, and fixations over 800ms were deleted. For analyses of the homograph, we deleted any trials in which subjects blinked during first-pass reading of the target homograph, resulting in 11% data loss.[5] Because the disambiguating region was a long, multi-word region, and trial exclusion based on blinks in the region resulted in the loss of a substantial

---

[5] Analyses completed without excluding any trials due to blinks on the homograph produced the same results, except that there was an additional effect of length on total time on the homograph ($p = .05$) and the effect of length on second pass time on the homograph pre-disambiguation became marginal ($p = .051$).

percentage of the data, we did not exclude trials based on blinks for analysis of the

disambiguating region. Mean fixation durations and regression probabilities by condition are

summarized in Table 2.

Linear mixed-effects models were fitted with the maximal random effects structure

justified by the design of the experiment, which included random item and participant intercepts

and slopes for sense, length, and their interaction. In order to fit the models, we used the lmer

function (glmer function to fit generalized linear mixed-effects regression models for binary

dependent variables) from the *lme4* package (version 1.1-10, Bates et al., 2015) within the R

Environment for Statistical Computing (R Development Core Team, 2015). We used sum coding

for the fixed effects of these predictors. Following Barr, Levy, Scheepers, & Tily (2013), we

assess significance of each predictor via a likelihood ratio test between a model with a fixed

effect of the predictor and one without it, maintaining identical random effects structure across

models.[6] Results of these models are summarized in Table 3.

Because the reordered access model predicts that there will not be an interaction of

meaning and length at disambiguation, the null hypothesis becomes theoretically important in

---

[6] Following Barr et al. (2013), when models failed to converge we removed correlation
parameters between random effects for analysis of that measure. When convergence issues
persisted, we either removed individual subjects with few observations or random slopes, always
maintaining at least random slopes for the fixed effect of interest in each test (e.g., random slopes
for length when comparing the maximal model to a model without an effect of length). In
Experiment 1, we removed two subjects who had observations in fewer than 20% of trials from
the analysis of first pass time on the homograph. For analysis of X←Y Regressions from
intermediate region 2 to the homograph, we removed the random slope of length for subjects
when testing for an effect of meaning, and the random slope of meaning for subjects and items
when testing for an effect of length. For analysis of X←Y Regressions from the disambiguation
region to the homograph, we removed the random slope of length for subjects and items when
testing for an effect of meaning, the random slopes of meaning and the interaction of meaning
and length for subjects and meaning for items when testing for an effect of length, and the
random slope of length for items when testing for an effect of the interaction.

this context (see Gallistel, 2009; Rouder, Morey, Speckman, & Province, 2012; Rouder, Speckman, Sun, Morey, & Iverson, 2009). Unfortunately, the traditional null hypothesis significance testing (NHST) framework, only allows us to reject the null if there is sufficient evidence in favor of the alternative (i.e., finding sufficient evidence for an interaction of meaning and length at p < .05), or fail to reject the null if there is not substantial evidence for the alternative (i.e., p > .05). However, failing to reject the null is not the same as finding support for the null. Indeed traditional NHST does not provide a framework for quantifying support for the null. Therefore, in addition to reporting the results of linear mixed-effects models, we also computed a Bayes factor for our critical measures. A Bayes factor is the ratio of the marginal likelihoods of the data D under two different statistical models, which allows us to directly assess the relative support for non-null (H_1) versus null (H_0) hypotheses (we represent Bayes factor in log base 10 space for transparency of interpretation):

$$\text{BF} = \log_{10} \frac{P(D|H_1)}{P(D|H_0)}$$

Values (in log base 10 space) near 0 indicate similar marginal likelihoods under two models, positive values indicate support for the non-null model, and negative values indicate support for the null model. When trying to determine the degree of support for a given model over another, it has been suggested that a Bayes factor in log base 10 space whose absolute value is greater than 0.5 should be interpreted as providing "substantial" evidence, greater than 1 as providing "strong" evidence, and greater than 2 as providing "decisive" evidence (Jeffreys 1961, Kass & Raftery, 1995).

We follow Abbott and Staub (2015; see also Rouder et al., 2012) in computing Bayes factors in an ANOVA design, but use the lmBF function in the R package, *BayesFactor* (Morey, Rouder, & Jamil, 2015), in order to include random item and participant slopes to conform to the

principle of using maximal random effects structure as implied by the design (Barr et al., 2013).

For by-subjects analysis of each dependent measure (corresponding to a traditional by-subjects

F_1 ANOVA), we compared the maximal model (where `Sub` is the random factor for subjects):

```
Response ~ (Meaning*Length*Sub)
```

to models without a main effect of meaning:

```
Response ~ (Meaning*Length*Sub – Meaning)
```

without a main effect of length:

```
Response ~ (Meaning*Length*Sub – Length)
```

or without the interaction of meaning and length:

```
Response ~ (Meaning*Length*Sub – Meaning:Length)
```

respectively. We also carried out corresponding item-mean analyses in which `Sub` in the above

formulae is replaced by `Item`. Critically, by comparing the maximal model to a model without

the interaction of meaning and length, we are able to measure whether the evidence favors a

model that includes an interaction of meaning and length, or a model without the interaction. The

Bayes Factor package assumes a Cauchy prior on effect size (*t* distribution, centered at zero), and

we set the scale parameter (reflecting the spread of the prior distribution) to 0.5.[7] The Bayes

factors obtained from direct comparisons between the maximal model and different reduced

---

[7] To ensure that the results we obtained in our Bayes analyses were not dependent on the specific priors that we selected, we also computed Bayes factors for two of our critical measures (second pass time on the homograph post-disambiguation and probability of making an X←Y regression from the disambiguating region to the homograph) with alternate priors by adjusting the scale of the Cauchy distribution by a factor of 2 in either direction (i.e., scale parameter set to 0.25 and 1). A scale parameter of 0.25 associates even more probability to very small effects, whereas a scale parameter of 1 associates more probability with larger effects. Relative to our normal scale parameter of 0.5, if we assume a small effect size, with the scale parameter set to 0.25 the Bayes factor would more strongly favor the non-null model, whereas with the scale parameter set to 1 the null model would be more strongly preferred. For both Experiments 1 and 2, critical results remained the same regardless of the spread of the Cauchy prior.

models are presented in Table 3. Results of by participant (F_1) and by item (F_2) ANOVAs paralleled the results we obtained with linear mixed effects models, but for transparency they are included in Appendix B. For reading time measures, Bayes Factor and traditional ANOVA analyses were conducted over participant/item means; for the binary measures (e.g., regression probability), these analyses were conducted over logit-transformed proportions (following Abbott & Staub, 2015; Barr, 2008).

Table 2

Mean Fixation Durations and Regression probabilities for Exp. 1

	Dominant		Subordinate	
	Short	Long	Short	Long
Pre-disambiguation				
Homograph				
first pass duration	243 (7.1)	266 (7.5)	251 (6.9)	256 (8.1)
second pass duration	59 (6.2)	71 (7.9)	51 (5.9)	72 (7)
rereading time	295 (16)	280 (22)	284 (18)	291 (16)
Lengthener				
first pass duration		601 (16)		637 (17)
second pass duration		43 (6.9)		31 (5)
Intermediate 2				
first pass duration	525 (15)	490 (13)	537 (16)	512 (15)
Homograph ← Intermediate 2				
regression probability	0.14 (0.02)	0.02 (0.01)	0.14 (0.02)	0.01 (0.01)
Intermediate super-region				
first pass regressions out	0.21 (0.02)	0.27 (0.02)	0.18 (0.02)	0.27 (0.02)
Homograph ← Intermediate super-region				
regression probability	0.20 (0.02)	0.25 (0.02)	0.18 (0.02)	0.25 (0.02)
Post-disambiguation				
Homograph				
second pass duration	52 (5.7)	49 (5.5)	74 (7.5)	70 (6.9)
rereading time	279 (14)	255 (16)	295 (19)	289 (16)
total time	345 (11)	352 (10)	358 (11)	363 (12)
Disambiguating Region				
first pass time (ms/char)	29.8 (0.87)	29.5 (0.85)	29.3 (0.86)	29.8 (0.85)
first pass regressions out	0.5 (0.02)	0.46 (0.02)	0.55 (0.02)	0.54 (0.02)
go-past time	1285 (50.5)	1395 (60.7)	1429 (56.5)	1501 (61.2)

Disambiguating Word				
first pass duration	254 (5.4)	264 (6.0)	253 (6.2)	264 (5.7)
total time	342 (10.7)	362 (11.9)	368 (11.2)	358 (11)
go-past time	356 (16.5)	371 (19.2)	416 (20.2)	398 (18.7)
Homograph ← Disambiguation				
regression probability	0.19 (0.02)	0.19 (0.02)	0.25 (0.02)	0.24 (0.02)

Note. Standard Errors appear in parentheses.

Pre-disambiguation. There were no significant effects of meaning or length on first pass times on the homograph, lengthener, or intermediate regions.[8] There was a significant effect of length in second pass time on the homograph pre-disambiguation and a significant effect of length on the probability of making a regression out of the intermediate super-region as well as the probability of making an X←Y regression from the intermediate super-region to the homograph, such that, prior to disambiguation, readers were more likely to make regressions to the homograph and had longer second pass reading times on the homograph in the long conditions. The Bayes factor analyses confirmed the results we obtained in the linear mixed effects models. For second pass time on the homograph, first pass regressions out of the intermediate super-region, and X←Y regression from the intermediate super-region to the homograph, Bayes factor analyses revealed that the models with an effect of length were favored over the reduced models without.

This effect of length on second-pass reading times for the homograph might seem to suggest that maintaining multiple word meanings from a homograph over longer periods of time without disambiguation is effortful. However, analysis of rereading time revealed no significant difference in the amount of time spent rereading the homograph across conditions. Taken together, these results suggest that the effect of length that we observed in second pass time was

---

[8] We did not test for an effect of length (or the interaction of meaning and length) for first pass time on the lengthening region, since this region only existed in the long conditions.

primarily driven by a tendency for readers to make more regressions into the homograph in the long conditions, rather than a tendency to spend significantly longer actually rereading the homograph following a regression. Furthermore, there was also a significant effect of length on the probability of making an X←Y regression from the second intermediate region to the homograph, such that readers were more likely to regress from intermediate region 2 to the homograph in the short conditions (i.e., they were more likely to regress from intermediate region 2 to the homograph when those two regions were often adjacent (the short conditions) than when lengthening material intervened).

No other effects of length were significant, and no effects of meaning or the interaction of meaning and length were observed across any measures prior to disambiguation. For all pre-disambiguation measures, Bayes factors computed for the maximal model compared to the model without an effect of meaning were all less than -0.36, and without an interaction of meaning and length were all less than -0.28, demonstrating that the reduced models were favored over (or at least preferred with approximately equal likelihood to) models with effects of meaning or the interaction of meaning and length.

Post-disambiguation. There was a main effect of meaning in second pass time on the homograph post disambiguation, such that readers had longer second pass time on the homograph when it was disambiguated to its subordinate meaning. The Bayes analyses for the maximal model compared to one without an effect of meaning confirmed that the maximal model was favored. Again, we computed a measure of pure rereading time that did not average in zeros when no second pass time occurred. Although this measure showed numerically longer rereading times following subordinate resolutions, the effect was not significant, demonstrating that the effect we observed in second pass time was primarily driven by a tendency for readers to

make more regressions into the homograph in the subordinate conditions, rather than a tendency to spend a significantly longer time actually rereading the homograph following a regression.

Additionally, there was a main effect of meaning in the probability of making an X←Y regression from the disambiguating region to the homograph, such that regression paths targeting the homograph were more likely following subordinate disambiguating material. The Bayes analyses confirmed this result; the maximal model with an effect of meaning was favored over one without for the probability of making an X←Y regression from the disambiguating region to the homograph. The maximal model with an effect of meaning was also favored over one without for the probability of making a regression out of the disambiguating region in general, though this effect was only marginal in the results of the linear mixed-effects models ($p = .06$). Post-hoc analyses of first pass time, go-past time, and total time on the disambiguating word revealed no significant effects of meaning, length, or the interaction of meaning and length, suggesting that disambiguation unfolded over time as participants read the disambiguating regions of our stimuli, rather than being driven by encountering a specific word.

No other effects of meaning were significant, and no effects of length or the interaction of meaning and length were observed across any measures after disambiguation. For all post-disambiguation measures, Bayes factors computed for the maximal model compared to a model without an effect of length were all less than -0.23, and the Bayes factors computed for the maximal model compared to a model without an interaction of meaning and length were all less than -0.26, demonstrating that the reduced models were favored over models with effects of length or the interaction of meaning and length—critically, providing support for the model with a null interaction of meaning and length.

Table 3. *Results of Linear Mixed-Effects Models and Bayes Analyses for Experiment 1. Significance of predictors meaning, length, and their interaction is assessed via likelihood ratio tests. Significant values appear in bold. Bayes Factors are presented in log base 10 and reflect the results from direct comparisons of the maximal model to the reduced models without an effect of meaning, length, or the interaction of meaning and length.*

	Meaning				Length				Meaning:length			
	LMM		Bayes Factor		LMM		Bayes Factor		LMM		Bayes Factor	
	χ^2	p	subject	item	χ^2	p	subject	item	χ^2	p	subject	item
Pre-disambiguation												
Homograph												
first pass duration	< .001	0.98	-0.72	-0.5	2.62	0.11	-0.35	-0.56	2.22	0.14	-0.37	0
second pass duration	0.18	0.67	-0.41	-0.5	**3.81**	**0.05**	0.84	0.19	0.40	0.53	-0.28	-0.44
rereading time	< .001	0.99			0.003	0.96			0.09	0.77		
Lengthener												
first pass duration	2.13	0.14										
second pass duration	1.46	0.23										
Intermediate 2												
first pass duration	0.80	0.37			3.25	0.07			0.71	0.40		
Homograph ← Intermediate 2												
regression probability	1.83	0.18			**25.52**	**<0.001**			0.16	0.69		
Intermediate super-region												
first pass regressions out	1.39	0.24	-0.68	-0.38	**8.38**	**0.004**	2.7	0.74	2.00	0.16	-0.46	-0.46

Homograph ← Intermediate- super-region												
regression probability	2.17	0.14	-0.62	-0.36	**7.43**	**0.006**	1.24	0.82	1.71	0.19	-0.46	-0.48
Post-disambiguation												
Homograph												
second pass duration	**5.11**	**0.02**	1.54	0.43	0.23	0.63	-0.58	-0.4	<.001	1	-0.34	-0.39
rereading time	2.07	0.15			0.75	0.39			0.54	0.46		
total time	0.98	0.32	-0.39	-0.31	0.47	0.49	-0.86	-0.3	< .01	0.97	-0.36	-0.26
Disambiguating Region												
first pass time (ms/char)	0.01	0.91			0.008	0.93			0.19	0.67		
first pass regressions out	3.45	0.06	0.54	0.26	0.95	0.33	-0.32	-0.23	0.48	0.49	-0.32	-0.36
go-past time	2.37	0.12			1.94	0.16			0.09	0.76		
Disambiguating Word												
first pass time	<0.01	0.95			3.5	0.06			<0.01	0.93		
total time	0.56	0.45			0.11	0.74			1.73	0.19		
go-past time	0.6	0.44			0.24	0.62			0.47	0.49		
Homograph ← Disambiguation												
regression probability	**7.54**	**0.006**	1.01	0.43	0.03	0.86	-0.75	-0.45	< .001	0.98	-0.62	-0.32

Discussion

In Experiment 1, we investigated whether readers attempt to maintain multiple meanings of a moderately biased homograph encountered in neutral context, and if so, whether this maintenance is subject to resource constraints. To do so, we tested for the presence of digging-in effects in the eye movement record as a function of increasing amounts of intervening sentence material before disambiguation. The capacity-constrained model of lexical ambiguity resolution predicted an interaction of meaning and length, such that processing of the subordinate resolution would be especially difficult when additional material intervened before disambiguation, whereas the reordered access model predicted that the subordinate resolution would be harder to process independent of the amount of intervening material.

Consistent with the predictions of both models, which assume that the subordinate meaning is less active or less readily available than the dominant meaning, we found that readers experienced difficulty disambiguating to the subordinate meaning of the homograph. Critically, this effect was not modulated by the amount of intervening material before disambiguation as the capacity-constrained model would predict—readers were more likely to regress to the homograph, and had longer second pass times on the homograph when it had been disambiguated to its subordinate meaning, regardless of the amount of text between the homograph and disambiguation. Bayes factor analyses confirmed that the model with a null interaction of meaning and length was favored over a model with a non-null interaction for all critical measures. This lack of an interaction of meaning and length is consistent with the predictions of the reordered access model and the results of Rayner and Frazier (1989), and lends no support to models specifying resource-constraints on multiple meaning maintenance, such as the capacity-constrained model, suggesting instead that readers can maintain multiple meanings

without significantly taxing cognitive resources, or that they are not attempting to maintain multiple meanings at all, instead selecting only one meaning to maintain.

Critically, the effects of meaning that we observed only arose at disambiguation. Prior to making any fixations in the disambiguating region of the text, we only observed effects of sentence length on the eye movement record, and no effects of the to-be-disambiguated meaning (as was expected, since all regions prior to the disambiguating region were identical across meanings). These effects of length are likely explicable in theoretically uninteresting ways. First, subjects were more likely to make regressions from the intermediate super-region to the homograph, and spent longer re-reading the homograph pre-disambiguation in the long conditions. This can most straightforwardly be explained as a result of the increased opportunities for regressions provided by the longer ambiguous region. Second, the probability of making a regression from the second intermediate region to the homograph pre-disambiguation was greater in the short conditions. In these conditions, there was usually no intervening material between the regression launch site and the homograph (the exception being the subset of sentences that included material in the first intermediate region), so it is plausible that these short regressions were just regressions between neighboring words (a very common type of regression, see Rayner, 2009 for a review) rather than the longer regression paths we observed post-disambiguation.

Since we found only main effects of length prior to disambiguation and of meaning at disambiguation, and no evidence of the interaction predicted by the capacity-constrained model, the results of this experiment are consistent with the reordered access model of lexical ambiguity resolution. Although these results are consistent with those reported by Rayner and Frazier (1989) for highly biased homographs, they stand in contrast to those of Miyake et al. (1994),

who reported lexical digging-in effects for moderately biased homographs. This is striking given the close similarity of the design of Experiment 1 to Miyake et al.'s Experiment 2; both used moderately biased homographs and had comparable length manipulations (the difference between Miyake et al.'s long and short conditions was 3-5 words compared to our 3-6 words). Aside from our use of eyetracking (compared to Miyake et al.'s self-paced reading), the key difference between the designs of the experiments was the choice of controls for the subordinate resolutions of the homographs: our Experiment 1 used dominant resolutions of the same homographs, while Miyake et al. used unambiguous semantic associates. Additionally, because we compared dominant and subordinate resolutions in Experiment 1, our disambiguating material was necessarily different. Since the critical interaction of meaning and length reported in Miyake et al. emerged only after disambiguation, it is important to be able to directly compare reading times in this region. In order to test whether these factors explained the contrasting results of the two experiments, we designed a second experiment using unambiguous controls, thereby more directly replicating Miyake et al.'s design. Again, the capacity-constrained model predicts an interaction of meaning (here subordinate v. unambiguous) and length, whereas the reordered access model predicts only a main effect of meaning at disambiguation.

Experiment 2

Methods

  Participants. An additional sixty native English speakers from the University of California, San Diego received course credit for their participation in the study. All participants had normal or corrected-to-normal vision.

  Apparatus. The apparatus was identical to Experiment 1.

Materials. Materials were adapted from Experiment 1 by replacing the dominant

conditions with unambiguous conditions: homographs in the dominant conditions were replaced

with unambiguous semantic associates of the homograph's subordinate sense (e.g. *speech* in the

case of *toast*). These semantic associates were roughly matched to the homograph's overall

word form frequency, but differed in length. Lexical frequencies (per 400 million) for all stimuli

were computed via log-transformed HAL frequency norms (Lund & Burgess, 1996) using the

English Lexicon Project (Balota et al., 2007). The homographs had an average word form log

frequency of 9.08 (range 6.94-11.79)[9], and the unambiguous semantic associates had an average

log frequency of 9.3 (range 6.69-12.05). Homographs and semantic associates were on average

4.75 and 5.31 characters long respectively. Critically, across all conditions, the disambiguating

regions were now identical and instantiated the homograph's subordinate resolution. This

facilitated comparison of reading measures in the disambiguating regions across all condition, as

the lexical content of the regions was identical. Sample stimuli appear in Table 4, and the full set

of stimuli is listed in Appendix A.

Procedure. The procedure was identical to Experiment 1.

Table 4

Sample Stimuli for Experiment 2

Meaning	Length	Intro	Homograph	Intermediate 1*	Lengthener *	Intermediate 2*	Disambiguation
Unambiguous	Short	The	speech	I made		was really	eloquent and well-delivered.
Unambiguous	Long	The	speech	I made	a few days ago	was really	eloquent and well-delivered.
Subordinate	Short	The	toast	I made		was really	eloquent and well-delivered.
Subordinate	Long	The	toast	I made	a few days ago	was really	eloquent and well-delivered.

Note. Asterisked regions are part of the Intermediate Super-region

Table 5

[9] Lexical frequency information for one homograph, "tip-off" was unavailable and was excluded
from these descriptive statistics.

Mean Fixation Durations and Regression Probabilities for Exp. 2

	Unambiguous		Subordinate	
	Short	Long	Short	Long
Pre-disambiguation				
Homograph/control				
first pass duration	234 (6.2)	225 (5.3)	230 (5.8)	247 (7.8)
second pass duration	36 (5.2)	52 (5.9)	41 (5.4)	67 (7.2)
rereading time	266 (22)	278 (15)	237 (19)	295 (18)
Lengthener				
first pass duration		614 (15)		636 (18)
second pass duration		30 (5.8)		52 (8)
Intermediate 2				
first pass duration	542 (17)	509 (14)	530 (14)	485 (12)
Homograph/control ← Intermediate 2				
regression probability	0.09 (0.01)	0.01 (0.005)	0.13 (0.02)	0.01 (0.01)
Intermediate super-region				
first pass regressions out	0.13 (0.02)	0.19 (0.02)	0.18 (0.02)	0.24 (0.02)
Homograph/control ← Intermediate super-region				
regression probability	0.14 (0.02)	0.18 (0.02)	0.17 (0.02)	0.22 (0.02)
Post-disambiguation				
Homograph/control				
second pass duration	44 (5.7)	38 (4.5)	64 (6.7)	55 (5.9)
rereading time	260 (20)	242 (11)	280 (16)	265 (14)
total time	303 (10)	306 (9.4)	326 (12)	352 (12)
Disambiguating Region				
first pass time (ms/char)	29.3 (0.71)	28.3 (0.68)	29.3 (0.76)	29.7 (0.74)
first pass regressions out	0.43 (0.02)	0.43 (0.02)	0.46 (0.02)	0.47 (0.02)
go-past time	1611 (43.1)	1599 (42.0)	1777 (42.8)	1788 (53.7)
Disambiguating Word				
fist past time	256 (5.0)	257 (4.8)	268 (5.6)	262 (5.3)
total time	330 (8.9)	326 (10.2)	371 (10.3)	343 (9.5)
go-past time	348 (14.7)	387 (22.1)	396 (18.5)	366 (16.8)
Homograph/control ← Disambiguation				
regression probability	0.18 (0.02)	0.16 (0.02)	0.22 (0.02)	0.21 (0.02)

Note. Standard errors appear in parentheses.

Results

Data pooling and exclusion criteria were identical to Experiment 1. For analyses of the

homograph, deletion of trials for blinks during first-pass reading of the target homograph

resulted in 7.8% data loss.[10] Participants were correct on an average of 93.8% of comprehension

questions. Mean fixation durations and regression probabilities by condition are summarized in

Table 5. Results of model comparisons and Bayes analyses are summarized in Table 6.[11] As

with Experiment 1, the ANOVA results paralleled the results we obtained with linear mixed

effects models, but for transparency they are reported in Appendix C.

      Pre-disambiguation. There were no significant effects of meaning or length on first

pass times on the homograph/control, lengthener, or intermediate regions.[12] There were

significant effects of length in second pass time on the homograph/control pre-disambiguation,

first pass regressions out of the intermediate super-region, and the probability of making an

X←Y regression from the intermediate super-region to the homograph/control, such that, prior to

disambiguation, readers were more likely to make regressions to the homograph/control and had

longer second pass times on the homograph/control in the long conditions. Confirming these

results, Bayes analyses favored the maximal model over the model without an effect of length for

second pass time on the homograph, first pass regressions out of the intermediate super-region,

[10] Again, analyses completed without excluding any trials due to blinks on the homograph
produced the same results, except that the effect of meaning on total viewing time on the
homograph was non-significant ($p = .11$)

[11] As with Experiment 1, when models failed to converge we removed correlation parameters
between random effects for analysis of that measure, but there were again a few models for
which convergence issues persisted. In Experiment 2, for analysis of X←Y Regressions from
intermediate region 2 to the homograph, we removed the random slope of length for subjects
when testing for an effect of meaning, and the random slope of meaning for subjects when
testing for effects of length and the interaction. For analysis of the intermediate super-region
regression out probability, we removed the random slope of length for subjects when testing for
an effect of meaning, and the random slope of meaning for subjects when testing for effects of
length and the interaction. For analysis of X←Y Regressions form the disambiguating region to
the homograph, we removed the random slope of length for subjects when testing for effects of
meaning and the interaction, and the random slope of meaning for subjects when testing for the
effect of length.

[12] As with Experiment 1, we did not test for an effect of length (or the interaction of meaning and
length) for first pass time on the lengthening region, since this region only existed in the long
conditions.

and the probability of making an X←Y regression from the intermediate super-region to the homograph/control.

As with Experiment 1, we also computed a measure of pure rereading time that did not average in zeros when no regression occurred. There was a marginal effect of length on rereading times, such that people spent numerically longer rereading the homograph/control in the long conditions, but again, this demonstrates that the effect we observed in second pass time was primarily driven by a tendency for readers to make more regressions into the homograph/control in the long conditions.

As in Experiment 1, there was also a significant effect of length on the probability of making an X←Y regression specifically from the second intermediate region to the homograph/control, such that readers were more likely to regress from intermediate region 2 to the homograph/control in the short conditions when those two regions were often adjacent.

Unlike Experiment 1, we found an effect of meaning in the likelihood of making a regression out of the intermediate super-region, with readers more likely to make a regression out of the intermediate super-region when it was preceded by the homograph than when it was preceded by an unambiguous control. Indeed, the Bayes analyses for the maximal model compared to a model without an effect of meaning, favored the maximal model. This result confirms that our eye movement measures are picking up meaningful correlates of processing difficulty due to disambiguation of lexical meaning.

No other effects of length or meaning were significant, and no interactions of meaning and length were observed across any measures prior to disambiguation. For all other pre-disambiguation measures, Bayes factors computed for the maximal model compared to the model without an effect of meaning were all less than -0.01, and without an interaction of

meaning and length were all less than -0.1, demonstrating that the reduced models were favored over (or at least preferred with approximately equal likelihood to) models with effects of meaning or the interaction of meaning and length.

Post-disambiguation. There were main effects of meaning in second pass time and total time on the homograph/control following disambiguation, such that readers had longer second pass times on the homograph once it was disambiguated to its subordinate meaning than they had on the unambiguous control word. The Bayes analyses confirmed that the maximal models were favored over models without effects of meaning for both second pass time and total time. We again computed a measure of pure rereading which patterned like second-pass time numerically, but was not significant. There was also a main effect of meaning on the probability of making an X←Y regression from the disambiguating region to the homograph/control, such that readers were more likely to make regressions from the disambiguating region to the subordinately disambiguated homograph than to the unambiguous control word. Indeed, the Bayes analyses for the probability of making an X←Y regression from the disambiguating region to the homograph/control confirmed that the maximal model was preferred over a model without an effect of meaning. Consistent with these effects, we also found a main effect of meaning in go-past time for the disambiguating region. Unlike Experiment 1, post-hoc analyses of the disambiguating word revealed a significant effect of meaning on the total time spent reading the disambiguating word, such that participants spent longer total time reading the disambiguating word in the subordinate condition than the unambiguous condition.

No other effects of meaning were observed and no effects of length or the interaction of meaning and length were observed following disambiguation. For all post-disambiguation measures, Bayes factors computed for the maximal model compared to a model without an effect

of length were all less than -0.08, and Bayes factors computed for the maximal model compared to a model without an interaction of meaning and length were all less than -0.1, demonstrating that the reduced models were favored over models with effects of length or the interaction of meaning and length—providing support for the model with a null interaction of meaning and length.

Table 6. *Results of Linear Mixed-Effects Models and Bayes Analyses for Experiment 2. Significance of predictors meaning, length, and their interaction is assessed via likelihood ratio tests. Significant values appear in bold. Bayes Factors are presented in log base 10 and reflect the results from direct comparisons of the maximal model to the reduced models without an effect of meaning, length, or the interaction of meaning and length.*

	Meaning				Length				Meaning:length			
	LMM		Bayes Factor		LMM		Bayes Factor		LMM		Bayes Factor	
	χ^2	p	sub	item	χ^2	p	sub	item	χ^2	p	sub	item
Pre-disambiguation												
Homograph/control												
first pass duration	1.85	0.17	-0.3	-0.27	0.83	0.36	-0.77	-0.44	3.49	0.06	-0.21	-0.1
second pass duration	2.14	0.14	-0.01	-0.11	**6.89**	**0.009**	2	0.68	0.49	0.48	-0.3	-0.3
rereading time	0.25	0.62			2.74	0.098			0.51	0.47		
Lengthener												
first pass duration	0.91	0.34										
second pass duration	3.14	0.08										
Intermediate 2												
first pass duration	15.46	0.28			19.52	0.11			14.59	0.33		
Homograph/control ← Intermediate 2												
regression probability	0.13	0.72			**32.82**	**<0.001**			1.54	0.21		
Intermediate super-region												
first pass regressions out	**6.60**	**0.01**	0.8	0.77	**4.55**	**0.03**	1.16	0.51	0.36	0.55	-0.48	-0.38
Homograph/control ← Intermediate super-region												
regression probability	3.17	0.08	0.31	0.15	**4.29**	**.04**	1.02	0.32	0.33	0.57	-0.38	-0.44
Post-disambiguation												
Homograph/control												

second pass duration	**8.23**	**0.004**	2.25	1.02	0.92	0.34	-0.35	-0.24	0.001	0.97	-0.37	-0.38
rereading time	2.25	0.13			0.94	0.33			0.006	0.94		
total time	**4.03**	**0.045**	1.5	0.45	1.43	0.23	-0.08	-0.3	0.98	0.32	-0.1	-0.34
Disambiguating Region												
first pass time (ms/char)	1.07	0.30			0.12	0.73			0.90	0.34		
first pass regressions out	2.21	0.14	-0.12	-0.04	0.04	0.83	-0.62	-0.36	0.04	0.85	-0.55	-0.3
go-past time	**12.54**	**< 0.001**			< 0.001	0.99			0.04	0.85		
Disambiguating Word												
first pass time	1.82	0.18			0.29	0.59			0.49	0.48		
total time	**10.43**	**0.001**			2.35	0.12			1.76	0.18		
go-past time	0.59	0.44			0.04	0.85			3.65	0.06		
Homograph/control ← Disambiguation												
regression probability	**4.64**	**0.03**	0.61	0.33	0.53	0.46	-0.22	-0.32	0.22	0.64	-0.31	-0.31

Discussion

In Experiment 2, we attempted a more direct replication of Miyake et al. (1994, Experiment 2). Following their design, we compared subordinate resolutions of moderately biased homographs to identical sentence frames with unambiguous controls, rather than the dominant resolutions of the same homographs as in Experiment 1. The results of Experiment 2 were parallel to those of Experiment 1 in all key respects. As in Experiment 1, prior to reaching the disambiguating region, eye movements exhibited effects of sentence length, which are not central to our current question and likely theoretically uninteresting (see our discussion of pre-disambiguation length effects for Experiment 1). Interestingly, an effect of meaning (ambiguous v. unambiguous) emerged prior to disambiguation that we did not observe in Experiment 1. Prior to reaching the disambiguating region, readers made more regressions out of the intermediate regions to reread the beginning of the sentence when it contained a homograph than when it contained an unambiguous control. Since we did not observe a difference in the pre-disambiguation regression rates between sentences containing dominant and subordinate homographs in Experiment 1, the difference in Experiment 2 likely reflects more effortful processing of an ambiguous word relative to an unambiguous word roughly matched on word form frequency (e.g., Sereno et al., 2006). Although initial processing of our homographs and unambiguous controls did not significantly differ (as is typical of highly-biased homographs, Simpson & Burgess, 1985), this difficulty in later measures (still prior to encountering any disambiguating information) suggests that processing difficulty for our moderately-biased homographs fell somewhere in between that of highly-biased and balanced homographs (for which initial processing is usually slowed, Rayner & Duffy, 1986).

After disambiguation, consistent with Experiment 1, we only observed effects of

meaning—readers spent longer total time reading the disambiguating word in the ambiguous conditions, and were more likely to regress to and spent longer second pass time on ambiguous homographs than unambiguous controls. These effects are again predicted by both the reordered access model and the capacity-constrained model, since, in both models, the subordinate meaning of a homograph is less readily available (i.e., less activated) than the single meaning of an unambiguous control. However, this effect of meaning was not modulated by the amount of intervening material prior to disambiguation as the capacity-constrained model would predict, and as Miyake et al. (1994) found. Indeed, as with Experiment 1, the Bayes factors computed for all critical post-disambiguation measures favored a model with a null interaction of meaning and length over one with a non-null interaction.

The fact that we again found a main effect of meaning at disambiguation that did not interact with length (as the capacity-constrained model would have predicted) lends further support to the reordered access model of lexical ambiguity resolution. Experiment 2 therefore rules out the possibility that different control conditions are responsible for the differences between our Experiment 1 results and those of Miyake et al. (1994, Experiment 2).

General Discussion

In two experiments, we investigated the processing of moderately biased homographs embedded in neutral preceding contexts. By varying the length of sentence material that intervened between the homograph and subsequent disambiguation, we sought to determine whether readers attempt to maintain multiple meanings of an ambiguous word presented without prior disambiguating information, and whether this meaning maintenance is subject to resource constraints. Consistent with both the reordered access model and the capacity-constrained model,

we found that disambiguating to the subordinate meaning was more difficult than disambiguating to the dominant meaning. In neither experiment did we find evidence for resource constraints on lexical ambiguity resolution: disambiguating to the subordinate meaning never become more difficult with increasing material. This second result runs counter to the predictions of the capacity-constrained model of lexical ambiguity resolution, and the previous results reported by Miyake et al. (1994). They found that increasing the distance between an ambiguous word and its disambiguation (thereby depleting cognitive resources, e.g., working memory), indeed made dispreferred resolutions especially difficult to process. The design of our experiments differed minimally from theirs, featuring moderately biased homographs, approximately the same additional distance to disambiguation between long and short sentence versions, and (in Experiment 2) the same choice of controls for subordinately-resolved homographs, namely unambiguous words.

The key remaining difference is the task itself: Miyake et al. used self-paced reading, while we used eyetracking. It is plausible that, given the generally lower resolution of self-paced reading (e.g., Witzel, Witzel, & Forster, 2012) and the fact that the crucial interaction was significant at $p < 0.05$ by subjects only, Miyake et al. observed a false positive. Indeed, careful inspection of their total reading time data following disambiguation (where they report their critical interaction of meaning and length) seems to show that the effect is being driven by effects of length when processing the unambiguous conditions. They report a difference in total reading time between the long and short ambiguous conditions of 22 ms (long =1737, short = 1715), and a difference between the long and short unambiguous conditions of -122 ms (long = 1552, short = 1654). While this is still an interaction of meaning and length at disambiguation, their capacity-constrained model would specifically predict an interaction driven by increased

reading time following disambiguation to the subordinate meaning in the long condition relative to the short condition, whereas they show a larger effect of length (that goes in the opposite direction) in the unambiguous conditions. Finally, most of their effects did not emerge immediately at disambiguation, but rather in spillover, and were pushed toward the end of the sentence, potentially further obscuring their results with sentence wrap-up effects.

While our results are inconsistent with the results of Miyake et al., they are consistent with the results of Rayner and Frazier (1989). They found that increasing the distance between an ambiguous word and its disambiguation had no effect on resolutions to the subordinate meaning—subordinate resolutions were more difficult than dominant resolutions, but the magnitude of this main effect of meaning did not vary as a function of length. They used highly biased homographs, for which, theoretically in the scope of the capacity-constrained model, initial activation of the subordinate meaning might have been so low that further effects of length could not be observed. However, the fact that we extended their results to moderately biased homographs demonstrates that their lack of an interaction was not likely due to floor effects in subordinate activation. Our results also go beyond those of Rayner and Frazier in using more tightly controlled disambiguating-region material (in Exp. 2), in drawing evidence from a wider range of eye movement measures, and in quantifying evidence in favor of the null hypothesis of no interaction between meaning and length by computing Bayes factors.

One could argue that perhaps our failure to find an interaction of meaning and length was due to the fact that even our short sentence versions were too long to show multiple-meaning maintenance. That is, perhaps the distance between the homographs and disambiguation in our short conditions (4.16 words on average) was not short enough to provide evidence for the maintenance of multiple meanings. We think this is unlikely given the converging results of

Rayner and Frazier (1989). In their short condition, the homograph was immediately followed by the disambiguating word and they still failed to find a difference between subordinate resolutions in that condition and their long condition, where 3-4 words intervened between the homograph and disambiguation. They argued that this suggested immediate resolution of lexical ambiguities even without contextual disambiguating information. Alternatively, one might instead question whether our length manipulation was simply too limited to detect any digging-in effects. Digging-in effects should manifest as positive correlations between the length of intervening material and any of our critical, post-disambiguation measures for subordinate resolutions (since more regressions and longer second pass time are indicative of more difficult processing). However, we find no evidence for a relationship between length of intervening material for a given item (measured as the length, in words, of the intermediate super-region) and any of our critical, post-disambiguation measures (all ps > .2).

Implications for models of lexical ambiguity resolution

Overall, then, the bulk of pertinent results on lexical disambiguation (our results and those of Rayner & Frazier, 1989) suggest one of two theoretical possibilities. First, readers may not attempt to maintain multiple meanings of an ambiguous word that they encounter in a neutral context, instead committing to one interpretation very rapidly—in the case of biased homographs, typically the dominant interpretation, as suggested by the reordered-access model. Under this explanation, since readers never attempt to maintain multiple meanings, whether or not cognitive resources are depleted during sentence comprehension, should have no effect on the reader's ultimate resolution—if they initially selected the correct meaning, reading will proceed easily, and if they initially selected the incorrect meaning, reading will likely be

disrupted, but the degree of disruption will not increase with more material. Second, readers may be able to maintain multiple meanings without significantly taxing available cognitive resources as suggested by the simplest probabilistic ranked-parallel models such as surprisal.

However, the idea that multiple meanings can be maintained without taxing cognitive resources may not be psychologically plausible. The number of possible interpretations of a sentence generally grows exponentially with its length (Church & Patil, 1982), and no known algorithm can exhaustively explore all possible interpretations in time linear in sentence length (which is the relationship between sentence length and processing time in humans: e.g., doubling a sentence's length approximately doubles the time required to read it). Recently, more cognitively realistic models of probabilistic sentence comprehension have been proposed which involve approximation algorithms intended to bring inferences close to the resource-unconstrained "ideal". Levy, Reali, and Griffiths (2009) proposed one such algorithm, the PARTICLE FILTER, in which limited resources are used to efficiently search the space of possible interpretations by repeated stochastic sampling as each incremental input word accrues. The stochasticity of the search gives rise to drift in the probabilities of alternative analyses, and the longer an ambiguity goes unresolved the more likely one of the interpretations is to be lost altogether. Thus, these more cognitively realistic models of probabilistic sentence comprehension predict that digging-in effects might arise during lexical ambiguity resolution, therefore making behavioral predictions analogous to the capacity-constrained model (see also Tabor & Hutchins, 2004 for a demonstration of how digging-in effects are predicted to arise from a related class of dynamical, self-organizing processing models). However, we found no evidence for digging-in effects in lexical ambiguity resolution, and therefore no evidence for the capacity-constrained model or particle filter.

Comparing lexical ambiguity and syntactic ambiguity

In principle, the processing of lexical ambiguity and syntactic ambiguity could well be fundamentally similar: both types of ambiguity might require the maintenance of multiple alternative representations, and multiple sources of ambiguity create a combinatorial explosion of overall possible interpretations that poses fundamental computational challenges. Arguments both for (e.g., MacDonald, 1993; MacDonald, Pearlmutter, & Seidenberg, 1994) and against (Traxler, Pickering & Clifton, 1998) this view have been advanced in the literature. The results from syntactic ambiguity resolution during the reading of garden-path sentences, suggest that readers are subject to resource constraints when resolving syntactic ambiguities, which gives rise to digging-in effects. For example, consider the reading of garden-path sentences as in (7) & (8):

(7) While the man hunted the deer ran into the woods.

(8) While the man hunted the deer that was brown and graceful ran into the woods

Although the ambiguity in these two sentences is structurally identical—in particular, the noun phrase containing *the deer* should be parsed as a new clause subject, not as the object of *hunted*—readers experience substantially more difficulty recovering from the ambiguity in (8) than in (7) (Ferreira & Henderson, 1991, 1993; Frazier & Rayner, 1982; Tabor & Hutchins, 2004).

If the processing of both types of ambiguity were fundamentally similar, then we would expect similar digging-in effects to emerge in lexical ambiguity resolution. However, that is not what we found. Our results are unlike those for syntactic ambiguity resolution, as we found no evidence for digging-in effects in lexical ambiguity resolution. These differences in how lexical

and syntactic ambiguities are managed pose challenges for accounts of ambiguity resolution that characterize syntactic ambiguity resolution purely as a type of lexical ambiguity resolution, suggesting instead that the two types of ambiguity may be represented differently, may impose different resource demands, and/or may be managed differently in human sentence comprehension.

Conclusion

Across two studies, we found no evidence for digging-in effects in lexical ambiguity resolution, and therefore no evidence for the capacity-constrained model. The ease (or difficulty in the case of subordinate disambiguation) with which readers disambiguate to each meaning did not increase with intervening material as the capacity-constrained model or the particle filter would have predicted. Instead, taken together with the results of Rayner and Frazier (1989), our results suggest that, in the absence of prior disambiguating context, either readers are able to maintain multiple meanings without significantly taxing cognitive recourses, or readers commit to one interpretation very rapidly—typically the more frequent interpretation—as suggested by the reordered-access model.

Acknowledgements

References

Abbott, M. J. & Staub, A. (2015). The effect of plausibility on eye movements in reading:
Testing E-Z Reader's null predictions. *Journal of Memory and Language, 85*, 76-87.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B.,…, Treiman,
R. (2007). The english lexicon project. *Behavior Research Methods, 39,* 445-459.

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

Barr, D. J., Levy, R., Scheepers,, C. & Tily, H. J. (2013). Random effects structure for
confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*,
255-278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Package 'lme4'

Church, K. & Patil, R. (1982). Coping with syntactic ambiguity, or how to put the block in the
box on the table. *Computational Linguistics, 8,* 139–149

Colbert-Getz, J. & Cook, A. E. (2013). Revisiting effects of contextual strength on the
subordinate bias effect: Evidence from eye movements. *Memory & Cognition, 41*, 1172-
1184.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in
reading. *Journal of Memory and Language*, 27, 429-446.

Ferreira, F. & Henderson, J. M. (1991). Recovery from misanalysis of garden-path sentences.
Journal of Memory and Language, 30, 725-745.

Ferreira, F. & Henderson, J. M. (1993). Reading processes during syntactic analysis and
reanalysis. *Canadian Journal of Experimental Psychology, 47*, 247-275.

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension:

Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178-210.

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review, 116*, 439-453.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In Proceedings of NAACL (Vol. 2, pp. 159-166).

Hudson, S.B. & Tanenhaus, M.K. (1984). Ambiguity resolution in the absence of contextual bias. *Proceedings of the Sixth Annual Cognitive Science Meetings*.

Jeffreys, H. (1961). Theory of probability (3rd ed.). New York: Oxford University Press.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science, 20*, 137–194.

Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: linguistic comprehension and production. In R. Bod, J. Hay, & S. Jannedy (Eds.), Probabilistic linguistics (pp. 39–95). MIT Press.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773-795.

Leinenger, M. & Rayner, K. (2013). Eye movements while reading biased homographs: Effects of prior encounter and biasing context on reducing the subordinate bias effect. *Journal of Cognitive Psychology, 25*, 665-681.

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

Levy, R. P., Reali, F., & Griffiths, T. L. (2009). modeling the effects of memory on human online sentence processing with particle filters. In *Advances in Neural Information Processing Systems*, pp. 937-944.

Lund, K. & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-

occurrence. *Behavior Research Methods, Instruments, & Computers, 28*, 203-208.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language, 32*, 692-715.

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

Miyake, A., Just, M. A., & Carpenter, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language, 33*, 175-202.

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes Factors for Common Designs. Version 0.9.12-2. http://bayesfactorpcl.r-forge.r-project.org/

Onifer, W. & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9*, 225-236.

Pacht, J. M., & Rayner, K. (1993). The processing of homophonic homographs during reading: Evidence from eye movement studies. *Journal of Psycholinguistic Research*, 22, 252-271.

R Development Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62, 1457-1506.

Rayner, K., Cook, A. E., Juhasz, B. J., & Frazier, L. (2006). Immediate disambiguating of

    lexically ambiguous words during reading: Evidence from eye movements.

    British Journal of Psychology, 97, 467-482.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading:

    Effects of word frequency, verb complexity, and lexical ambiguity. *Memory &*

    Cognition, 14, 191-201.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous

    words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*,

    15, 779-790.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors

    for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356-374.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayeseian t-tests

    for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review, 16*,

    225-237.

Sereno, S. C., O'Donnell, P. J., & Rayner, K. (2006). Eye movements and lexical

    ambiguity resolution: Investigating the subordinate-bias effect. *Journal of*

    Experimental Psychology: Human Perception and Performance, 32, 335-350.

Sheridan, H., & Reingold, E. M. (2012). The time course of contextual influences during lexical

    ambiguity resolution: Evidence from distributional analyses of fixation durations.

    Memory & Cognition, *40*, 1122-1131.

Sheridan, H., Reingold, E. M., & Daneman, M. (2009). Using puns to study contextual

    influences on lexical ambiguity resolution: Evidence from eye movements. *Psychonomic*

    Bulletin & Review, 16, 875-881.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical
ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20*, 120-136.

Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the
recognition of ambiguous words. *Journal of Experimental Psychology: Human
Perception & Performance*, 11, 28-39.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of
context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645-659.

Tabor, W. & Hutchins, S. (2004). Evidence for self-organized sentence processing: Diggin-in
effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 431-
450.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in
the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and
Verbal Behavior, 18*, 427-440.

Traxler, M. J., Pickering M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical
ambiguity resolution. *Journal of Memory and Language, 39*, 558-592.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative
meaning frequency for 566 homographs. *Memory & Cognition, 22*, 111-126.

Wiley, J., & Rayner, K. (2000). Effects of titles on the processing of text and lexically
ambiguous words: Evidence from eye movements. *Memory & Cognition*, 28,
1011-1021.

Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye-
tracking, moving-window, and maze. *Journal of Psycholinguistic Research, 41*, 105-128.

Stimuli (Target Words are Underlined, Disambiguating Material appears in Italics. Sentence versions a & b are dominant disambiguations, versions c & d are subordinate disambiguations, and e & f show unambiguous controls. Experiment 1 included sentence versions a-d, Experiment 2 included sentence versions c-f)

1   a   The <u>nail</u> was obviously from Lisa's *chair because its armrest was loose.*
    b   The <u>nail</u> that John noticed was obviously from Lisa's *chair because its armrest was loose.*
    c   The <u>nail</u> was obviously from Lisa's *finger because it was pink.*
    d   The <u>nail</u> that John noticed was obviously from Lisa's *finger because it was pink.*
    e   The <u>jewelry</u> was obviously from Lisa's *finger because it was pink.*
    f   The <u>jewelry</u> that John noticed was obviously from Lisa's *finger because it was pink.*

2   a   The <u>buzz</u> last Friday was *coming from the broken alarm system.*
    b   The <u>buzz</u> in the hallways last Friday was *coming from the broken alarm system.*
    c   The <u>buzz</u> last Friday was *about the prom king and queen.*
    d   The <u>buzz</u> in the hallways last Friday was *about the prom king and queen.*
    e   The <u>gossip</u> last Friday was *about the prom king and queen.*
    f   The <u>gossip</u> in the hallways last Friday was *about the prom king and queen.*

3   a   The <u>deed</u> described by Andy was *an unmatched feat of heroism.*
    b   The <u>deed</u> described in great detail by Andy was *an unmatched feat of heroism.*
    c   The <u>deed</u> described by Andy was *for the house he inherited.*
    d   The <u>deed</u> described in great detail by Andy was *for the house he inherited.*
    e   The <u>lamp</u> described by Andy was *for the house he inherited.*
    f   The <u>lamp</u> described in great detail by Andy was *for the house he inherited.*

4   a   The <u>spade</u> got lost after the *gardener completed his work.*
    b   The <u>spade</u> that John was looking for got lost after the *gardener completed his work.*
    c   The <u>spade</u> got lost after the *card game the night before.*
    d   The <u>spade</u> that John was looking for got lost after the *card game the night before.*
    e   The <u>box</u> got lost after the *card game the night before.*

f   The <u>box</u> that John was looking for got lost after the *card game the night before.*

5   a   The highly respected <u>major</u> was actually still *classified as a hard science.*
   b   The highly respected <u>major</u> was, it turned out, actually still *classified as a hard science.*
   c   The highly respected <u>major</u> was actually still *in the lieutenant colonel's office.*
   d   The highly respected <u>major</u> was, it turned out, actually still *in the lieutenant colonel's office.*
   e   The highly respected <u>officer</u> was actually still *in the lieutenant colonel's office.*
   f   The highly respected <u>officer</u> was, it turned out, actually still *in the lieutenant colonel's office.*

6   a   The <u>pit</u> was actually really easy to *climb out of on one side.*
   b   The <u>pit</u>, she was happy to discover, was actually really easy to *climb out of on one side.*
   c   The <u>pit</u> was actually really easy to *cut out of the fruit she was eating.*
   d   The <u>pit</u>, she was happy to discover, was actually really easy to *cut out of the fruit she was eating.*
   e   The <u>stem</u> was actually really easy to *cut out of the fruit she was eating.*
   f   The <u>stem</u>, she was happy to discover, was actually really easy to *cut out of the fruit she was eating.*

7   a   The <u>tip-off</u> marked the beginning of the *criminal investigation by the district attorney's office.*
   b   The <u>tip-off</u> that everyone had been anticipating marked the beginning of the *criminal investigation by the district attorney's office.*
   c   The <u>tip-off</u> marked the beginning of the *final round of the basketball series.*
   d   The <u>tip-off</u> that everyone had been anticipating marked the beginning of the *final round of the basketball series.*
   e   The <u>fireworks</u> marked the beginning of the *final round of the basketball series.*
   f   The <u>fireworks</u> that everyone had been anticipating marked the beginning of the *final round of the basketball series.*

8   a   A <u>horn</u> can signal *danger to other drivers on the road.*
   b   A <u>horn</u> can be used to signal *danger to other drivers on the road.*
   c   A <u>horn</u> can signal *an animal's dominance over potential competitors.*
   d   A <u>horn</u> can be used to signal *an animal's dominance over potential competitors.*
   e   A <u>roar</u> can signal *an animal's dominance over potential competitors.*
   f   A <u>roar</u> can be used to signal *an animal's dominance over potential competitors.*

9   a   A <u>yard</u> is a really *fun place for a child to play.*

b A <u>yard</u> is by all accounts a really *fun place for a child to play.*
c A <u>yard</u> is a really *long distance for a baby to walk.*
d A <u>yard</u> is by all accounts a really *long distance for a baby to walk.*
e A <u>mile</u> is a really *long distance for a baby to walk.*
f A <u>mile</u> is by all accounts a really *long distance for a baby to walk.*

10 a A good <u>ruler</u> is essential to have in a *drawing class of any kind.*
    b A good <u>ruler</u>, everyone generally agrees, is essential to have in a *drawing class of any kind.*
    c A good <u>ruler</u> is essential to have in a *country that is at war.*
    d A good <u>ruler</u>, everyone generally agrees, is essential to have in a *country that is at war.*
    e A good <u>king</u> is essential to have in a *country that is at war.*
    f A good <u>king</u>, everyone generally agrees, is essential to have in a *country that is at war.*

11 a The <u>pot</u> they bought was the wrong kind for *cooking the exotic dish.*
    b The <u>pot</u> they bought and were very excited to try was the wrong kind for *cooking the exotic dish.*
    c The <u>pot</u> they bought was the wrong kind for *smoking immediately before bed.*
    d The <u>pot</u> they bought and were very excited to try was the wrong kind for *smoking immediately before bed.*
    e The <u>marijuana</u> they bought was the wrong kind for *smoking immediately before bed.*
    f The <u>marijuana</u> they bought and were very excited to try was the wrong kind for *smoking immediately before bed.*

12 a The <u>palm</u> was decorated with beautiful *henna tattoos in black and brown.*
    b The <u>palm</u> in the photograph was decorated with beautiful *henna tattoos in black and brown.*
    c The <u>palm</u> was decorated with beautiful *lights for the upcoming holiday season.*
    d The <u>palm</u> in the photograph was decorated with beautiful *lights for the upcoming holiday season.*
    e The <u>tree</u> was decorated with beautiful *lights for the upcoming holiday season.*
    f The <u>tree</u> in the photograph was decorated with beautiful *lights for the upcoming holiday season.*

13 a The <u>calf</u> is classified with *wild bulls in ancient traditions of idol-worship.*
    b The <u>calf</u>, I recently learned, is classified with *wild bulls in ancient traditions of idol-worship.*
    c The <u>calf</u> is classified with *two other muscles that flex the ankle.*
    d The <u>calf</u>, I recently learned, is classified with *two other muscles that flex the ankle.*

e The <u>leg</u> is classified with *two other muscles that flex the ankle.*

f The <u>leg</u>, I recently learned, is classified with *two other muscles that flex the ankle.*

14 a The <u>term</u> his colleague advised him not to *serve in office led to his impeachment.*

b The <u>term</u> his colleague tried harder than ever to advise him not to *serve in office led to his impeachment.*

c The <u>term</u> his colleague advised him not to *use was the controversial word "terrorism."*

d The <u>term</u> his colleague tried harder than ever to advise him not to *use was the controversial word "terrorism."*

e The <u>word</u> his colleague advised him not to *use was the controversial word "terrorism."*

f The <u>word</u> his colleague tried harder than ever to advise him not to *use was the controversial word "terrorism."*

15 a The <u>star</u> was too hard to see through the *telescope on that cloudy night.*

b The <u>star</u> that Michele was hoping to find was too hard to see through the *telescope on that cloudy night.*

c The <u>star</u> was too hard to see through the *crowd of fans and photographers.*

d The <u>star</u> that Michele was hoping to find was too hard to see through the *crowd of fans and photographers.*

e The <u>actress</u> was too hard to see through the *crowd of fans and photographers.*

f The <u>actress</u> that Michele was hoping to find was too hard to see through the *crowd of fans and photographers.*

16 a The middle <u>digit</u> is usually the *number that's lowest in area code prefixes.*

b The middle <u>digit</u> is, in the US, usually the *number that's lowest in area code prefixes.*

c The middle <u>digit</u> is usually the *finger that's used in obscene gestures.*

d The middle <u>digit</u> is, in the US, usually the *finger that's used in obscene gestures.*

e The middle <u>finger</u> is usually the *finger that's used in obscene gestures.*

f The middle <u>finger</u> is, in the US, usually the *finger that's used in obscene gestures.*

17 a His mother's <u>trust</u> was something he hoped *to earn back through his good behavior toward her.*

b His mother's <u>trust</u> was very important to him and something he hoped *to earn back through his good behavior toward her.*

c His mother's <u>trust</u> was something he hoped *would contain enough funds for his college education.*

d His mother's <u>trust</u> was very important to him and something he hoped *would contain enough funds for his college education.*

e His mother's gift was something he hoped *would contain enough funds for his college education.*

f His mother's <u>gift</u> was very important to him and something he hoped *would contain enough funds for his college education.*

18 a The <u>mold</u> was probably not *growing behind the wall.*

b The <u>mold</u> he was looking for was probably not *growing behind the wall.*
c The <u>mold</u> was probably not *of a famous actress's face.*
d The <u>mold</u> he was looking for was probably not *of a famous actress's face.*
e The <u>sculpture</u> was probably not *of a famous actress's face.*
f The <u>sculpture</u> he was looking for was probably not *of a famous actress's face.*

19 a I thought the <u>finish</u> was great for a *rookie racecar driver in her first race.*
    b I thought the <u>finish</u> we were discussing was great for a *rookie racecar driver in her first race.*
    c I thought the <u>finish</u> was great for a *cedar bookshelf in the living room.*
    d I thought the <u>finish</u> we were discussing was great for a *cedar bookshelf in the living room.*
    e I thought the <u>paint</u> was great for a *cedar bookshelf in the living room.*
    f I thought the <u>paint</u> we were discussing was great for a *cedar bookshelf in the living room.*

20 a Sam's first <u>letter</u> was most likely *a thank-you message to his grandma.*
    b Sam's first <u>letter</u> when he was young was most likely *a thank-you message to his grandma.*
    c Sam's first <u>letter</u> was most likely *an upper-case A he learned in preschool.*
    d Sam's first <u>letter</u> when he was young was most likely *an upper-case A he learned in preschool.*
    e Sam's first <u>vowel</u> was most likely *an upper-case A he learned in preschool.*
    f Sam's first <u>vowel</u> when he was young was most likely *an upper-case A he learned in preschool.*

21 a Kenji noticed that one <u>temple</u> was inexplicably *offering no Passover service at all.*
    b Kenji noticed that one <u>temple</u>, much to his surprise, was inexplicably *offering no Passover service at all.*
    c Kenji noticed that one <u>temple</u> was inexplicably *swollen in the patient's forehead.*
    d Kenji noticed that one <u>temple</u>, much to his surprise, was inexplicably *swollen in the patient's forehead.*
    e Kenji noticed that one <u>brow</u> was inexplicably *swollen in the patient's forehead.*
    f Kenji noticed that one <u>brow</u>, much to his surprise, was inexplicably *swollen in the patient's forehead.*

22 a The <u>speaker</u> would be louder if *her microphone were on.*
    b The <u>speaker</u> that they were listening to would be louder if *her microphone were on.*
    c The <u>speaker</u> would be louder if *its amplifier were working properly.*
    d The <u>speaker</u> that they were listening to would be louder if *its amplifier were working properly.*

e   The <u>stereo</u> would be louder if *its amplifier were working properly.*

f   The <u>stereo</u> that they were listening to would be louder if *its amplifier were working properly.*

23   a   The <u>affair</u> was supposed to be kept a *secret from the lovers' spouses.*

      b   The <u>affair</u> was, according to most sources, supposed to be kept a *secret from the lovers' spouses.*

      c   The <u>affair</u> was supposed to be kept a *strictly black tie event.*

      d   The <u>affair</u> was, according to most sources, supposed to be kept a *strictly black tie event.*

      e   The <u>event</u> was supposed to be kept a *strictly black tie event.*

      f   The <u>event</u> was, according to most sources, supposed to be kept a *strictly black tie event.*

24   a   The <u>crook</u> is the most interesting part of a *crime novel or detective story.*

      b   The <u>crook</u> is, in my opinion anyway, generally the most interesting part of a *crime novel or detective story.*

      c   The <u>crook</u> is the most interesting part of a *shepherd's staff in medieval paintings.*

      d   The <u>crook</u> is, in my opinion anyway, generally the most interesting part of a *shepherd's staff in medieval paintings.*

      e   The <u>hook</u> is the most interesting part of a *shepherd's staff in medieval paintings.*

      f   The <u>hook</u> is, in my opinion anyway, generally the most interesting part of a *shepherd's staff in medieval paintings.*

25   a   The <u>toast</u> I made was really *delicious with blackberry jam.*

      b   The <u>toast</u> I made a few days ago was really *delicious with blackberry jam.*

      c   The <u>toast</u> I made was really *eloquent and well-delivered.*

      d   The <u>toast</u> I made a few days ago was really *eloquent and well-delivered.*

      e   The <u>speech</u> I made was really *eloquent and well-delivered.*

      f   The <u>speech</u> I made a few days ago was really *eloquent and well-delivered.*

26   a   Except for a small <u>wave</u>, it seemed that the *lake was completely calm that afternoon.*

      b   Except for a small <u>wave</u> that Rob barely noticed, it seemed that the *lake was completely calm that afternoon.*

      c   Except for a small <u>wave</u>, it seemed that the *shy girl had ignored him completely.*

      d   Except for a small <u>wave</u> that Rob barely noticed, it seemed that the *shy girl had ignored him completely.*

      e   Except for a small <u>smile</u>, it seemed that the *shy girl had ignored him completely.*

      f   Except for a small <u>smile</u> that Rob barely noticed, it seemed that the *shy girl had ignored him completely.*

27   a   When she broke the <u>glass</u>, I was annoyed that the *wine spilled all over the carpet.*

b   When she broke the <u>glass</u> by accident the other day, I was annoyed that the *wine spilled all over the carpet.*

c   When she broke the <u>glass</u>, I was annoyed that the *screen door would need to be replaced.*

d   When she broke the <u>glass</u> by accident the other day, I was annoyed that the *screen door would need to be replaced.*

e   When she broke the <u>handle</u>, I was annoyed that the *screen door would need to be replaced.*

f   When she broke the <u>handle</u> by accident the other day, I was annoyed that the *screen door would need to be replaced.*

28   a   The <u>pupil</u> was worrying Susan because *he was misbehaving again.*

b   The <u>pupil</u> was worrying Susan even more that day because *he was misbehaving again.*

c   The <u>pupil</u> was worrying Susan because *it was dilated again.*

d   The <u>pupil</u> was worrying Susan even more that day because *it was dilated again.*

e   The <u>eye</u> was worrying Susan because *it was dilated again.*

f   The <u>eye</u> was worrying Susan even more that day because *it was dilated again.*

29   a   The <u>chest</u> was determined to be the *muscle group most often toned in the gym.*

b   The <u>chest</u> was, according to what I heard, determined to be the *muscle group most often toned in the gym.*

c   The <u>chest</u> was determined to be the shipwreck's most valuable artifact by the expert.

d   The <u>chest</u> was, according to what I heard, determined to be the *shipwreck's most valuable artifact by the expert.*

e   The <u>treasure</u> was determined to be the shipwreck's most valuable artifact by the expert.

f   The <u>treasure</u> was, according to what I heard, determined to be the *shipwreck's most valuable artifact by the expert.*

30   a   The <u>bulb</u> finally ended up *burning out and needing replacement.*

b   The <u>bulb</u> we'd forgotten about finally ended up *burning out and needing replacement.*

c   The <u>bulb</u> finally ended up *sprouting in the flowerbed outside.*

d   The <u>bulb</u> we'd forgotten about finally ended up *sprouting in the flowerbed outside.*

e   The <u>flower</u> finally ended up *sprouting in the flowerbed outside.*

f   The <u>flower</u> we'd forgotten about finally ended up *sprouting in the flowerbed outside.*

31   a   Krista thought the <u>note</u> was supposed to be *an apology from her boyfriend.*

b   Krista thought the <u>note</u> she was telling me about was supposed to be *an apology from her boyfriend.*

c   Krista thought the <u>note</u> was supposed to be *a B-flat on the clarinet.*

d   Krista thought the <u>note</u> she was telling me about was supposed to be *a B-flat on the clarinet.*

e   Krista thought the <u>noise</u> was supposed to be *a B-flat on the clarinet.*

f   Krista thought the <u>noise</u> she was telling me about was supposed to be *a B-flat on the clarinet.*

32   a   The <u>tissue</u> was simply too *rough on his nose so he bought a better kind.*

b   The <u>tissue</u>, he quickly realized, was simply too *rough on his nose so he bought a better kind.*

c   The <u>tissue</u> was simply too *damaged from the surgery and needed to be replaced.*

d   The <u>tissue</u>, he quickly realized, was simply too *damaged from the surgery and needed to be replaced.*

e   The <u>skin</u> was simply too *damaged from the surgery and needed to be replaced.*

f   The <u>skin</u>, he quickly realized, was simply too *damaged from the surgery and needed to be replaced.*

Experiment 1 Analyses of Variance with random error by subjects (F_1) and by items (F_2) for the critical dependent measures on the target word and post target regions.

Variable	$F_1(1,59)^*$	p	$F_2(1,31)$	p
First Pass (Hom)				
Length	1.74	0.19	**6.31**	**0.02**
Meaning	0.06	0.81	0.01	0.92
Length:Meaning	1.17	0.28	2.0	0.17
Second Pass pre-disambiguation (Hom)				
Length	**8.3**	**< 0.01**	**4.16**	**0.05**
Meaning	0.36	0.55	0.19	0.67
Length:Meaning	0.46	0.50	0.69	0.41
First Pass Regressions Out (Interm Super Region)				
Length	**19.86**	**< 0.001**	**9.12**	**< 0.01**
Meaning	0.48	0.49	0.55	0.46
Length:Meaning	0.67	0.42	0.57	0.46
Regressions Hom ← Interm Super Region				
Length	**10.71**	**< 0.01**	**8.25**	**< 0.01**
Meaning	0.37	0.55	0.88	0.36
Length:Meaning	0.4	0.53	0.39	0.54
Second Pass post-disambiguation (Hom)				
Length	0.35	0.56	0.31	0.58
Meaning	**12.23**	**< 0.001**	**5.54**	**0.03**
Length:Meaning	< .001	0.997	< .001	0.997
Total time (Hom)				
Length	0.09	0.76	1.27	0.27
Meaning	1.26	0.27	1.23	0.28
Length:Meaning	0.04	0.85	0.01	0.92
First Pass Regressions Out (Disambig Region)				
Length	1.34	0.25	1.86	0.18
Meaning	**5.76**	**0.02**	**4.39**	**0.04**
Length:Meaning	0.89	0.35	0.71	0.41
Regressions Hom ← Disambig Region				
Length	0.02	0.9	0.09	0.76
Meaning	**10.4**	**< 0.01**	**5.44**	**0.03**
Length:Meaning	0.01	0.94	0.1	0.75

Note. The degrees of freedom for the F_1 analysis of first pass duration on the homograph were actually (1,56) due to the exclusion of 3 subjects for missing data in at least 1 condition. Values in bold are significant at $\alpha = .05$.

Appendix C
Experiment 2 Analyses of Variance with random error by subjects (F_1) and by items (F_2)
for the critical dependent measures on the target word and post target regions.

Variable	$F_1(1,59)$*	p	$F_2(1,31)$	p
First Pass (Hom)				
Length	0.27	0.61	1.14	0.29
Meaning	2.08	0.16	1.37	0.25
Length:Meaning	2.76	0.1	2.89	0.1
Second Pass pre-disambiguation (Hom)				
Length	**15.21**	**< 0.001**	**7.07**	**0.01**
Meaning	3.25	0.08	2.57	0.12
Length:Meaning	1.03	0.31	0.51	0.48
First Pass Regressions Out (Interm Super Region)				
Length	**9.77**	**< .01**	**6.58**	**0.02**
Meaning	**7.41**	**< 0.01**	**7.98**	**< 0.01**
Length:Meaning	0.05	0.83	0.32	0.58
Regressions Hom ← Interm Super Region				
Length	**9.31**	**< 0.01**	**5.59**	**0.02**
Meaning	**5.18**	**0.03**	3.96	0.06
Length:Meaning	0.01	0.91	0.34	0.56
Second Pass post-disambiguation (Hom)				
Length	1.5	0.23	0.91	0.35
Meaning	**16.12**	**< 0.001**	**8.79**	**< 0.01**
Length:Meaning	0.02	0.90	0.04	0.84
Total time (Hom)				
Length	2.92	0.09	1.17	0.29
Meaning	**12.61**	**< 0.001**	**5.56**	**0.02**
Length:Meaning	2.56	0.12	0.62	0.44
First Pass Regressions Out (Disambig Region)				
Length	0.14	0.71	0.11	0.74
Meaning	3.6	0.06	2.16	0.15
Length:Meaning	0.2	0.66	< 0.01	0.95
Regressions Hom ← Disambig Region				
Length	2.08	0.16	0.63	0.43
Meaning	**7.3**	**< 0.01**	**4.99**	**0.03**
Length:Meaning	0.16	0.69	< 0.01	0.94

Note. The degrees of freedom for the F_1 analysis of first pass duration on the homograph were actually (1,58), due to the exclusion of 1 subject for missing data in at least 1 condition. Values in bold are significant at $\alpha = .05$.