Psychology and Aging
2014, Vol. 29, No. 2, 213–228

0882-7974/14/$12.00 http://dx.doi.org/10.1037/a0036282

Adult Age Differences in Wrap-Up During Sentence Comprehension: Evidence From Ex-Gaussian Distributional Analyses of Reading Time

Brennan R. Payne and Elizabeth A. L. Stine-Morrow
University of Illinois at Urbana-Champaign

We report a secondary data analysis investigating age differences in the effects of clause and sentence wrap-up on reading time distributions during sentence comprehension. Residual word-by-word self-paced reading times were fit to the ex-Gaussian distribution to examine age differences in the effects of clause and sentence wrap-up on both the location and shape of participants' reaction time (RT) distributions. The ex-Gaussian distribution showed good fit to the data in both younger and older adults. Sentence wrap-up increased the central tendency, the variability, and the tail of the distribution, and these effects were exaggerated among the old. In contrast, clause wrap-up influenced the tail of the distribution only, and did so differentially for older adults. Effects were confirmed via nonparametric vincentile plots. Individual differences in visual acuity, working memory, speed of processing, and verbal ability were differentially related to ex-Gaussian parameters reflecting wrap-up effects on underlying reading time distributions. These findings argue against simple pause mechanisms to explain end-of-clause and end-of-sentence reading time patterns; rather, the findings are consistent with a cognitively effortful view of wrap-up and suggest that age and individual differences in attentional allocation to semantic integration during reading, as revealed by RT distribution analyses, play an important role in sentence understanding.

Keywords: aging, ex-Gaussian, individual differences, RT distributional analysis, sentence processing, working memory, wrap-up

Aging is associated with divergent trajectories of change in language understanding. Some aspects of language processing are spared or even show improvements with advancing age (e.g., visual word recognition, verbal ability; Lien et al., 2006; Verhaeghen, 2003). However, effortful processes related to comprehension and memory for message-level semantics in spoken and written language show considerable age-related decline (Dagerman, MacDonald, & Harm, 2006; Federmeier, 2007; Johnson, 2003; Payne et al., 2014; Stine-Morrow, Miller, & Hertzog, 2006; Stine-Morrow & Miller, 2009). One mechanism that has often been implicated in the encoding and maintenance of message-level semantics during sentence comprehension is the so-called *wrap-up effect*, a phenomenon marked by relative increases in processing time at clause and sentence boundaries (Just & Carpenter, 1980; Rayner, Kambe, & Duffy, 2000). Though this is an empirically robust phenomenon (for a review, see Payne & Stine-Morrow, 2012), the mechanisms underlying these effects are not well understood.

Wrap-Up Effects in Sentence Comprehension

The observation of unique processing episodes at the ends of clauses and sentences has been of interest in psycholinguistics for at least 50 years (Aaronson & Scarborough, 1976; Caplan, 1972; Fodor, & Bever, 1965; Garret, Bever, & Fodor, 1966; Jarvella, 1971; Kimball, 1973). The phrase wrap-up was coined by Just and Carpenter (1980), who described this phenomenon as

> [a] special computational episode [that] occurs when a reader reaches the end of a sentence . . . . The processes that occur during sentence wrap-up involve a search for referents that have not been assigned, the construction of interclause relations (with the aid of inferences, if necessary), and an attempt to handle any inconsistencies that could not be resolved within the sentence. (p. 345)

Payne and Stine-Morrow (2012) have referred to such explanations of wrap-up as semantic integration theories, in that they predict that these characteristic peaks in reading time reflect a cognitively demanding segmentation process, in which message-level semantic information is integrated across clauses and sentences to enable a coherent and stable representation of the text in memory (Kintsch, 1998; Millis & Just, 1994; Rayner et al., 2000; Stine-Morrow, Millinder, Pullara, & Herman, 2001; Stine-Morrow & Miller, 2009). Semantic integration theories are supported by studies showing that clause and sentence wrap-up are dependent upon ongoing cognitive workload during sentence comprehension. For example, wrap-up has been shown to be moderated by prop-

Brennan R. Payne and Elizabeth A. L. Stine-Morrow, The Beckman Institute for Advanced Science and Technology, and the Department of Educational Psychology, University of Illinois at Urbana-Champaign.

We thank Daniel Spieler and two anonymous reviewers for their helpful discussions and comments on earlier drafts of this article. This research was supported by the National Institute on Aging (R01 AG013935). The first author was supported by a National Institutes of Health Training Grant (T32-HD055272) and a Beckman Institute Predoctoral Fellowship during the preparation of this article.

Correspondence concerning this article should be addressed to Brennan R. Payne, Cognitive Science of Teaching and Learning Division, Department of Educational Psychology, University of Illinois at Urbana-Champaign, 226 Education Building, 1310 South Sixth Street, Champaign, IL 61820-6990. E-mail: payne12@illinois.edu

erties of the text, including manipulations of domain knowledge (Miller & Stine-Morrow, 1998; Sharkey & Sharkey, 1987; Wiley & Rayner, 2000), conceptual complexity (Haberlandt & Graesser, 1989; Stine-Morrow et al., 2010), gap-filling (Balogh et al., 1998), and syntactic and semantic ambiguity resolution (Hagoort, 2003; Luo, Yan, & Zhou, 2013; Payne et al., 2014), as well as properties of the individual, including individual differences in verbal ability and linguistic experience (Payne et al., 2012; Stine-Morrow et al., 2008), comprehension goals (Fallon, Peelle, & Wingfield, 2006; Stine-Morrow et al., 2006), and maintenance of concurrent working memory load (Smiler, Gagne, & Stine-Morrow, 2003).

In contrast to this view of wrap-up as a semantic integration process, some have argued that wrap-up effects may reflect an automatic and obligatory process that is driven by low-level perceptual characteristics of the text (e.g., commas and periods; Hill & Murray, 2000) or the monitoring of implicit prosodic contour alone (Hirotani, Frazier, & Rayner, 2006). Hirotani et al. (2006) argued that because much of semantic and syntactic analysis is incremental (Rayner & Clifton, 2009; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), it is not clear what incomplete work is left to be done at the ends of clauses and sentences. They showed that wrap-up effects occur during reading at points corresponding to intonational phrase boundaries in speech—points that would not be expected to correspond to high cognitive load (e.g., vocatives such as *John*, *go to the library for me*). They proposed a dwell-time account to explain such findings, in which the increased processing time at clause boundaries reflects an automatic dwell response that "need not be related to the amount of work to be done at the clause boundary" (p. 426).

On these accounts, wrap-up is an obligatory process that is not driven by ongoing processing difficulty during sentence understanding. Consistent with this claim are findings by Hirotani et al. (2006), who showed that manipulations of the surface length of a preceding clause did not moderate the magnitude of wrap-up, and Warren, White, and Reichle (2009), who showed that manipulations of syntactic complexity did not moderate the magnitude of wrap-up. These findings suggest that wrap-up effects may be fixed to some extent, occurring at points of both high and low demand within a sentence. Notably, such models predict that clause and sentence boundaries will trigger an increase in reading time overall, but that these effects will be invariant across properties of the text and individual, reflecting a uniform delay in processing time before moving forward in the text (see Hill & Murray, 2000; Hirotani et al., 2006 for a discussion).

Aging and Individual Differences in Wrap-Up

Individual differences in the magnitude of wrap-up effects have been shown to be predictive of comprehension and sentence memory in both younger and older adults (Miller & Stine-Morrow, 1998; Schroeder, 2011; Stine-Morrow et al., 2001). Importantly, a number of studies have converged on the finding that deficits in comprehension and memory are magnified among older adults who fail to allocate attention at clause and sentence boundaries (Smiler et al., 2003; Stine, 1990; see Stine-Morrow et al., 2006; Stine-Morrow & Miller, 2009 for reviews). Such findings implicate wrap-up as a potential compensatory mechanism that older adults may be able to take advantage of during online sentence processing to maintain optimal comprehension in the face of declines (Stine-Morrow et al., 2010).

Payne and Stine-Morrow (2012) investigated age differences in the cognitive workload associated with clause and sentence-wrap up, using a gaze-contingent boundary-change eye-tracking paradigm. They examined effects of wrap-up on parafoveal word processing on the word immediately after a clause or sentence boundary, with the logic that if wrap-up is cognitively demanding, it should reduce the amount of parafoveal processing on the word following wrap-up (cf. Henderson & Ferreria, 1990; White, Rayner, and Liversedge, 2005; see also White, Warren, & Reichle, 2011). They found substantial age differences in the effects of sentence wrap-up on parafoveal word processing, but only for later-pass measures that included trials with refixations and regressions. These findings suggested that wrap-up effects did not induce a high degree of cognitive workload across all trials, but rather were driven by a subset of trials that were the most demanding, and thus where more time was needed to process the target word (see also Rayner, Castelhano, & Yang, 2009, 2010). Importantly, these findings were not consistent with obligatory accounts of wrap-up: on the trial with the longest fixations (presumably the most demanding trials), age-differences in the costs of wrap-up were largest.

Distributional Analyses of RT

Without exception, the studies reviewed above on wrap-up effects in reading have focused on changes in central tendency by comparing mean differences in reading or fixation times. Although the mean provides important information, using measures of central tendency alone to examine condition differences in RT has a number of drawbacks. The most significant concern with relying only on the mean to draw conclusions about underlying processing mechanisms is that RT distributions are never normally distributed and the shape can vary as a function of condition differences. Extreme values are often trimmed based on some criterion, but to the extent that variables of interest differentially affect the central tendency, spread, and tail of the distribution, basing conclusions on the mean of the distribution from which values at the tail are eliminated can provided a distorted picture of the phenomenon of interest.

A number of researchers have advocated investigating performance beyond the mean using distributional analytic methods (Balota et al., 2008; Balota & Yap, 2011). Although these methods have existed in the literature for some time (Luce, 1986; Ratcliff, 1979; van Zandt, 2000), distributional analyses of RTs have recently gained increasing traction, being applied to attention (Castel, Balota, Hutchison, Logan, & Yap, 2007; Tse, Balota, Yap, Duchek, & McCabe, 2010), episodic memory (Rohrer & Wixted, 1994; Criss, 2010), visual word recognition (Balota et al., 2008), and sentence processing (Staub, 2011; Staub & Benatar, 2013) domains. Typically, distributional analyses involve fitting a probability distribution that closely approximates the characteristic unimodal and skewed distribution of RTs to empirical RTs. Though there are a number of potential probability distributions for RT data, the ex-Gaussian distribution, which is a convolution of the Gaussian (normal) distribution and the exponential distribution, typically provides good fit to RT data while using a small number of parameters (Heathcote, Brown, & Cousineau, 2004;

Heathcote, Brown, & Mewhort, 2002; van Zandt, 2000; Balota & Yap, 2011).

In the ex-Gaussian distribution, RT distributions are summarized by three parameters. Two parameters are from the Gaussian component of the distribution: central tendency (i.e., the location of the distribution) is reflected in μ, and variability in the modal portion of the distribution is captured by σ, the standard deviation. The rate parameter, τ, from the exponential component of the ex-Gaussian distribution, reflects the degree of rightward slowing in the tail of the distribution. Importantly, conditions that impact μ result in a shift in the distribution, conditions that impact σ result in increased variance in RTs, and conditions that impact τ selectively increase the subset of trials that are the slowest.

Recently, a number of studies have applied these distributional analyses to decompose RTs and eye-fixation times during language processing tasks (including sentence reading), showing that linguistic variables impact multiple components of an RT distribution (e.g., Staub, White, Drieghe, Hollway, & Rayner, 2010; White & Staub, 2012). Importantly, reading and fixation times during sentence processing show a characteristic rightward skew, like most RT data. This "slow tail" is ignored in traditional analyses, despite the fact that different sorts of linguistic complexity may differentially affect not only the location of a distribution, but its underlying shape as well. For example, Staub et al. (2010) showed that word frequency influences both the location (μ) and the slow tail (τ) of the distribution of eye-fixations in sentence reading (see Balota & Spieler, 1999 for similar findings in lexical decision tasks), signifying that frequency effects occurred on almost all trials, thus shifting the distribution, but that word frequency effects became larger as fixation times became slower. These findings suggest that frequency effects are much stronger on the most demanding trials (see Balota & Spieler, 1999 for an extensive discussion of similar effects in lexical decision time tasks). Similarly, Johnson, Staub, and Fleri (2012) recently showed, in a naming task, that transposed letter effects influence the τ parameter only, such that on the slowest trials, the presence of a transposed letter neighbor inhibited naming performance. They argued that this effect on the tail might be attributable to a mechanism whereby lexical access occasionally fails on a subset of trials, rather than a reliable effect of transposition on access. It is unlikely that evidence for such a mechanism would be found simply by comparing group means alone, because the underlying shape of the distribution is hidden once the data are averaged.

Balota and Yap (2011) argued that there are three general approaches to examining the effects of variables on underlying RT distributions. These include (a) fitting empirical distributions (such as the ex-Gaussian) to generate a small number of parameters that descriptively summarize RT distributions, (b) adopting a nonparametric approach by plotting percentiles of RT distributions across different conditions, and lastly (c) using a computationally explicit process model that links psychological processes to specific RT distributional characteristics. Indeed, Ratcliff's (1979; Ratcliff & McKoon, 2008) diffusion model has been highly successful in modeling two-alternative force choice RT data from paradigms where a RT and paired binary decision are available for all observations (e.g., lexical decision data). Although this model has been successful in modeling decision phenomenon in a number of cognitive tasks, it is limited in its usefulness for modeling online word-by-word sentence-processing data, where RT data are available for all words within a sentence without any directly observable binary decision criteria for each RT, and ultimately, where accuracy is not binary (Caplan & Waters, 2013; Ferreira & Patson, 2007).

It is important to note that one must be cautious in directly interpreting derived ex-Gaussian parameters, given that they cannot be directly linked to specific cognitive mechanisms, such as attentional lapses or changes in response criteria (Matzke & Wagenmakers, 2009). Nevertheless, ex-Gaussian models do provide a very useful descriptive method to understand not only the influence of variables on RTs above and beyond changes in the mean, but also individual differences in RT distributions (Schmiedek, Oberauer, Wilhelm, Süss, & Witmann, 2007; Yap et al., 2012), which would be likely obscured with methods that focus on the mean alone.

For instance, RT distributional analyses have provided valuable insight into aging and individual differences in attentional control (Schmiedek et al., 2007; Jackson, Balota, Duchek, & Head, 2012; Tse et al., 2010). This is particularly true for the τ component, which has been shown to share substantial variance with individual differences in age, speed of processing, working memory, risk for Alzheimer's disease, and white matter integrity in healthy older adults and early Alzheimer's patients (Jackson et al., 2012; Schmiedek et al., 2007; Tse et al., 2010). Such results are similar to findings demonstrating increased age-related intraindividual variability across a number of cognitive tasks (Hultsch et al., 2008; Salthouse & Berish, 2005).

The link between τ in attentional control tasks and individual differences in WM and processing speed is consistent with the worse performance rule (Coyle, 2003), whereby the slowest RTs are most highly related to performance on tasks of fluid cognitive ability. Some have argued that condition differences that impact τ partially reflect failures of continuous attentional control throughout task performance (Jackson et al., 2012; McVay & Kane, 2012; Schmiedek et al., 2007; Tse et al., 2010; West, 2001), such that individuals who have difficulty maintaining goals and suppressing irrelevant information produce especially slow RTs on a subset of trials, resulting in a pronounced rightward skew in the RT distribution (see Coyle, 2003 for a discussion). Schmeidek and colleagues (2007), using structural equation modeling, found a latent correlation of −.90 between WM span and τ components derived from attention tasks, indicating that individuals with lower working memory spans had a larger proportion of extreme RTs in tasks involving strong response competition demands. Importantly, distributional parameters derived from performance in attentional control (Balota et al., 2008), visual word recognition (Yap & Balota, 2007; Yap et al., 2012), and sentence reading (Staub & Benatar, 2013) tasks have shown substantial test–retest reliability, suggesting that these parameters can be viewed as stable individual differences.

Collectively, these findings suggest that, although parameters of RT distributions have not been directly linked through the use of explicit process models (as in two-choice decision tasks; Ratcliff & McKoon, 2008), certain distributional parameters (such as τ) may be more sensitive to breakdowns in executive control processes in normal aging (Jackson et al., 2012). However, individual differences in ex-Gaussian parameters have been primarily examined in attention and lexical performance tasks, and, to date, we know of no investigation of the effects of aging and individual differences in cognition on RT distributions in sentence processing.

The Current Study

The data reported in the current study are drawn from a large-scale individual difference study in which participants read single sentences in a self-paced word-by-word moving window paradigm, originally reported in Stine-Morrow et al. (2008). We utilize RT distributional analyses to address three questions.

First, do reading times at clause and sentence boundary sites differ from medial sites in the location (μ), scaling (σ), and slow tail (τ) of the distribution? Importantly, under an automatic/oculomotor explanation of wrap-up, clause and sentence wrap-up would be expected to impact a majority of trials in a uniform way (e.g., a shift in the location of the distribution). For example, Warren et al. (2009), based on computational modeling of wrap-up effects in the EZ-Reader model (Pollatsek, Reichle, & Rayner, 2006), argued that the best fitting models implicated an early low-level (likely oculomotor) mechanism that is responsive to the presence of phrase boundaries and punctuation, but is not impacted by variability in item difficulty. Under this account, boundary sites would be expected to trigger uniform wrap-up processes in reading, resulting in a shifting of the RT distributions at clause and sentence boundaries.

However, this account does not predict that one would see increased variability in RT distributions at boundary sites, nor would one expect to see a larger proportion of extreme RTs at clause or sentence boundaries. Similar (though not completely analogous) arguments have been made in the visual word recognition and attention literature, using distributional analyses to test automatic "head-start" models of lexical access (Balota et al., 2008), or "criterion shift" models of selective attention (Spieler et al., 2000), by examining the degree to which certain variables shift RT distributions versus impacting the slow tail of RT distributions. Note that the mechanisms proposed in the literature on two-choice decisions are only imperfect analogs to the processes ongoing during word-by-word reading time, given the lack of a clear "one-shot" process leading to a binary decision at every trial (i.e., every word), which is clearly present in tasks such as lexical decision, but is absent in naturalistic reading for comprehension.

In contrast, a semantic integration view, in which wrap-up varies depending on properties of the text and the individual (Payne et al., 2012; Stine-Morrow et al., 2001), predicts not only a shift in the RT distribution at clause and sentence boundaries, but that changes in the variance and slow tail of the distribution may also occur across individuals, reflecting differential responsiveness to processing demands and processing disruptions during wrap-up. Staub and Benatar (2013) presented a theory of eye fixation time distributions during sentence processing whereby variables that are found to have a stronger effect in the tail of the distribution index an increased proportion of trials in which normal processing is disrupted, distinct from processing difficulty that uniformly shifts the RT distribution (see Coyle, 2003; Schmiedek et al., 2007 for similar arguments in the attention literature). Under a semantic integration view, a larger proportion of extreme reading times may be expected at clause and sentence boundaries, indexing the disruption caused by the increased cognitive workload of interclause binding and integration processes in wrap-up.

Our second question focused on whether age differences emerge in the shape of reading time distributions at clause and sentence boundaries. Under a dwell-time account, distributional shifting at clause and sentence boundaries should be similar for younger and older adults. If global age-related slowing is implicated, however, this may simply result in a larger shift in the distribution (effect on μ) for older adults relative to the young, but no change in the underlying shape of the distribution. Under a semantic integration account however, there may be selectively strong wrap-up effects for the slowest trials, reflecting that wrap-up effects may induce a high cognitive demand on a subset of trials, an effect that may be larger for older adults, consistent with studies comparing mean age differences across various eye-movement measures (Payne & Stine-Morrow, 2012; Rayner et al., 2010).

Our third and final question was whether distributional parameters are related to individual differences in sensory and cognitive ability. If age differences in the slow tail of the distribution were found, this may implicate a mechanism associated with increased cognitive workload. However, to rule out that such findings are not attributable to other causes, such as a more conservative processing criteria among the old (Starns & Ratcliff, 2010), we also aimed to test whether age differences could be explained in terms of individual differences in sensory and cognitive ability.

If wrap-up effects are more demanding for older adults, one would expect distributional parameters in the tail of the distribution to be related to individual differences in fluid cognitive ability, as in the attentional control literature (Schmiedek et al., 2007; Tse et al., 2010). Indeed, many studies focusing on age differences in language comprehension have proposed that psychomotor speed and working memory play a substantial role in sentence comprehension (Caplan et al., 2011; Dagerman et al., 2006; DeDe et al., 2004; Kemper & Liu, 2007; Stine-Morrow et al., 2001; Payne et al., 2012, 2014), though there is much controversy surrounding the exact role these play in online interpretive processes (see Caplan & Waters, 1999; Just & Varma, 2002; MacDonald & Christiansen, 2002). We predicted that measures of working memory and speed would play a larger role in wrap-up effects on the tail of the distribution, consistent with a semantic integration view of wrap-up, in which such effects reflect increased cognitive demand (Payne & Stine-Morrow, 2012).

Moreover, given the increased interest in linguistic experience as an important individual difference in language processing (Payne et al., 2012, 2014; MacDonald & Christiansen, 2002), including in RT distributional studies of word recognition (Yap et al., 2012), we also examined whether verbal ability was related to RT distributional effects of clause and sentence wrap-up. Yap et al. (2012) showed that higher levels of vocabulary knowledge were related to facilitated effects on μ, σ, and τ in naming and lexical decision tasks. Similarly, higher levels of verbal ability have been shown to be predictive of facilitated lexical processing at the sentence level, in studies of mean reading time performance (Payne et al., 2012). However, Payne et al. (2012) also found that older adults with higher verbal ability showed *larger* effects of clause wrap-up (see also Yamani et al., 2012). However, it is yet unknown how these effects at the mean translate to changes in underlying RT distributions.

Lastly, sensory acuity plays an important role in language comprehension in older adults (Burke & Shafto, 2008 for a review), and is often considerably correlated with individual differences in cognitive performance (Anstey, Hofer, & Luszcz, 2003; Lindenberger & Baltes, 1994). However, sensory abilities are rarely examined in conjunction with cognitive individual differences in studies of online sentence processing. Recognition of marked clause and sentence boundaries requires fine visual discrimination, such that potential age differences in reading time distributions and effects of wrap-up could be driven in part by sensory ability, apart from age-related changes in cognition. Thus, in order to tease apart influences of sensory and cognitive ability on RT distributions during reading, we also examined the role of individual differences in visual acuity.

Method

Data Set

Stine-Morrow et al. (2008) provides an in depth description of participant characteristics, procedures, and measures. In the following, we review important characteristics of the data set relevant for the current study.

Participants

One hundred twenty-eight young adults (18–39 years) and 117 older and middle-aged adults (58–87 years) participated in the current study. One older adult was removed from the analysis because, after trimming excessively long trials (see below), the ex-Gaussian distribution could not be fit in the sentence final condition for that participant because of a lack of valid RT observations. Thus, analyses are based on 244 participants. Table 1 presents demographic and individual difference characteristics separately for the younger and older adult groups.

Measures

Participants completed a series of individual difference measures in a single battery. *Working memory span* (α = .73) was assessed by three tasks, the reading span and listening span tasks described in Stine and Hindman (1994) and the computation span task (Salthouse & Babcock, 1991). *Speed of processing* (α = .69) was assessed with two tasks, the letter comparison and pattern comparison tasks (Salthouse & Babcock, 1991). *Verbal ability* (α = .93) was measured with the ETS advanced vocabulary and extended range vocabulary tasks (Ekstrom, French, & Harman, 1976). *Visual acuity* was assessed with the Snellen chart, which was administered separately for the right and left eyes. The derived measure was the average Snellen ratio for both eyes.

Table 1
Demographics and Individual Differences

	Young (n = 117)		Old (n = 127)	
Measure	*M*	*SE*	*M*	*SE*
Demographics				
1. Age	20	.27	71	.66
2. Education	13	.13	17	.22
3. Self-reported health	1.67	.05	1.88	.08
4. Visual acuity	1.04	.03	.68	.03
Verbal ability				
5. Extended range	15.42	.60	32.36	.78
6. Advanced vocabulary	11.99	.37	25.29	.61
Working memory				
7. Reading span	4.73	.12	3.95	.12
8. Listening span	5.27	.10	3.91	.09
9. Computation span	6.44	.10	4.66	.20
Speed of processing				
10. Letter comparison	23.07	.43	18.41	.40
11. Pattern comparison	42.30	.69	30.81	.53

Self-Paced Reading Task

Participants read a series of 48 18-word sentences describing facts in nature, science and history (a total of 864 words per participant; e.g., *Since George Washington's time, United States Presidents have received presents from both ordinary citizens and heads of state*), in a noncumulative word-by-word moving window self-paced reading paradigm (Aaronson & Ferres, 1984). These items were first used in Stine-Morrow et al. (2001). Each sentence was followed by a filler sentence to ensure that sentence final reading times were not contaminated by retrieval planning or task preparation. Filler sentences were not analyzed. The self-paced sentence reading task was followed by two tasks involving reading larger expository and narrative passages, but those data were not used in the current study (see Stine-Morrow et al., 2008), given that our primary focus was on sentence processing. Each of the 864 words was coded as either marking a minor clause boundary (n = 157), a sentence boundary (n = 48), or a sentence medial word, which did not mark a sentence or clause boundary (n = 659). The total number of observations was (864 $\times$ 244 =) 210, 816.

Data Analysis

Reading times were trimmed for outliers within person and within sentence position (medial, clause-final, sentence-final) at the 5th standard deviation. This is a conservative approach, removing only extreme values that were attributable to external distraction, failure to follow directions, and other factors unrelated to on-task processes. Following this, a three-step analytical approach was conducted to fit the ex-Gaussian distribution to each participant's data.

First, a residual analysis was conducted to control for a number of item-level predictors of word reading time that covary with relative sentence position. Following Hofmeister (2011) and Jaeger (2010), a linear mixed effects model was fit to the RT data for all items and all participants, with the following covariates: (a) word length (in syllables), (b) log of the word frequency, (c) word class (open vs. closed), (d) Yngve depth (a measure of syntactic depth based on the cumulative number of incompletely parsed phrase-structure rules),[1] (e) linear word position in the sentence, and (f) trial order, as well as a random intercept for subjects. Note

[1] Yngve depth (1960) is a metric derived from computational linguistics that has been commonly used to estimate syntactic complexity in studies of language comprehension and production (Cheung & Kemper, 1992; Stine-Morrow, Loveless, & Soederberg, 1996). In brief, each test sentence was constructed into a parsing tree representing its syntactic structure and branches were consecutively numbered from right to left (beginning with 0). The Yngve depth was coded for each word as the sum of all branches leading from the terminal node to the top of the tree.

that this is a conservative approach, as some of the variance being statistically controlled by these variables is likely driving wrap-up processes (e.g., changes in syntactic depth across embedded clauses). Nevertheless, this approach allows for us to test for the unique influence of wrap-up, after partialing out variance in other common item-level predictors of RT (cf. Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). All covariates were significant predictors of word reading time. Item-level residuals from this model were then extracted, representing variability in reading times independent of the covariate regressors. Each subject's residual reading times in each of the three word-position conditions was then fit to an ex-Gaussian distribution using quantile maximum likelihood estimation with the program QMPE 2.18 (Heathcote et al., 2004).

Because residual reading times contain negative values, and the domain of the exponential component of the ex-Gaussian distribution is non-negative, a constant value was added to all residual reading times in every condition for every subject. Before adjustment, the minimum residual RT was −1,103 ms (mean residual RT = 0). A constant of 1,200 ms was added to all residual RTs, setting the minimum RT equal to approximately 100 ms, so that all values submitted for analyses were non-negative and did not approach zero (mean adjusted Residual RT = 1,200 ms, min = 96 ms, max = 5,006 ms). Note that adding a constant essentially shifts the entire residual RT distribution for all individuals and all conditions. Thus, this adjustment had no influence on group or condition differences in estimates of μ (though it does increase the total value of μ estimates across the board by 1,200 ms) and no effects on the estimates for σ or τ.

Finally, to formally test for differences in reading time distributions as a function of age group and boundary site, the extracted parameters from the ex-Gaussian distribution were submitted to separate statistical analyses. Linear mixed effects models were fit separately to each of the three parameters. Models were estimated with fixed effects for age group, sentence position, and an age group by sentence position interaction. A random intercept across subjects was also included so that these were equivalent to hierarchical linear models (Quené & van den Bergh, 2004; Traxler, 2009).[2] Because sentence position has three levels, this variable was parameterized with the medial condition as the reference group, forming two dummy coded variables: the clause wrap-up (CW) effect (medial vs. clause boundary) and the sentence wrap-up (SW) effect (medial vs. sentence boundary). When age was a significant moderator (in two-way interactions), these comparisons were explored in follow-up analyses by fitting age-separate models. Markov chain Monte Carlo sampling was used to generate 95% highest posterior density intervals (95% HDI) for all fixed effect parameters (Baayen et al., 2008; Kruschke, 2011). Models were estimated using the lmer function in the lme4 package (Bates, Maechler, & Bolker, 2013) in R (R Core Team, 2013).

Results

Age Differences in Mean Reading Time

Before presenting the ex-Gaussian analyses, effects are presented at the mean level for comparison. Table 2 presents means for the unresidualized raw RTs and the adjusted residual RTs for medial, clause-final, and sentence-final words, separately for younger and older adults. Results of a linear mixed effects analysis on the adjusted residual reading times (with crossed random effects for subjects and items) revealed a reliable Age × CW interaction, $\gamma = 61$; 95% HDI: [55, 62], and a reliable Age × SW interaction, $\gamma = 243$; 95% HDI: [231, 254]. Older adults showed a reliable clause wrap-up effect of 45 ms (95% HDI: [27, 62]) and a reliable sentence wrap-up effect of 378 ms (95% HDI: [347, 409]). However, younger adults' clause final residual reading times were actually 18ms faster than for medial words (after adjusting for item-level covariates; 95% HDI: [−31, −6]), though they did show a sentence-final wrap-up effect of 134 ms (95% HDI: [113, 154]).

Table 2
Raw Mean RTs and Adjusted Residual RTs by Age Group and Sentence Position (in ms)

RT	Young		Old		Total	
	M	*SE*	*M*	*SE*	*M*	*SE*
Raw RTs						
Medial	440	20	599	31	516	19
Clause final	494	25	711	44	595	26
Sentence final	653	34	1036	62	821	36
Adjusted residual RTs						
Medial	1,197	18	1,177	26	1,188	16
Clause final	1,177	22	1,218	38	1,196	21
Sentence final	1,330	30	1,551	55	1,427	31

Ex-Gaussian Model Fitting

The ex-Gaussian models converged normally for each participant. The modeled distribution provided excellent fit to the RT data for each subject in each condition. At the item level, a pseudo R^2 statistic (squared sample correlation between observed and predicted values; see Singer & Willet, 2003) was calculated separately for each participant in each condition. The average pseudo R^2 value was .95 (SD = .03; range = .891, .999) across subjects and conditions, indicating excellent fit. This can be observed graphically in the Figure in the Appendix, which plots observed item RTs against ex-Gaussian model predicted item RTs for a randomly selected set of 10 participants from each of the three sentence position conditions. In addition, a visual summary of the degree of fit of the ex-Gaussian models can be seen in Figure 3 by comparing the model estimated vincentiles against the empirical vincentiles (see below for a discussion of the vincentile plot). Ex-Gaussian model fits did not differ for younger or older adults.

Age Differences in Reading Time Distributions

Figure 1 presents the mean values for each of the ex-Gaussian parameters in each condition and age group, along with 95% confidence intervals. Results from the linear mixed-effects models are reported below, separately for each parameter.

For the μ parameter (Figure 1, top panel), the Age × SW interaction was reliable, $\gamma = 201$; 95% HDI: [144, 255], but the

[2] Note that there is no random effect for items because the ex-Gaussian parameters are computed across items separately for each subject.

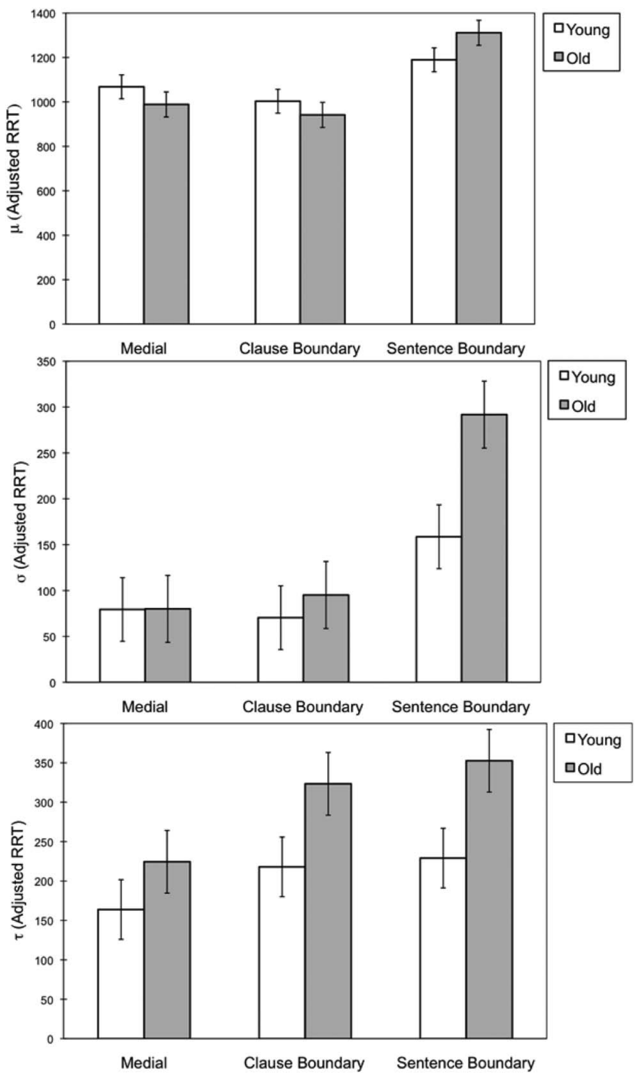


Figure 1. Means and standard errors of ex-Gaussian parameters by age group and word position. Values are based on residual reading times that were rescaled by adding a constant to remove nonzero values for ex-Gaussian model fitting (see text for details).

Age × CW interaction was not, γ = 18; 95% HDI: [−37, 73]. The source of the Age × SW interaction was that the difference in μ between medial and sentence-final conditions was larger for the older adults, γ = 322; 95% HDI: [273, 372], compared with the young, though the SW effect was still different from zero in the young adult group, γ = 121; 95% HDI: [93, 149]. Collapsing across age groups, the CW effect was reliable and negative, γ = −56; 95% HDI: [−86, −29], indicating that distribution in the clause-final condition was actually shifted slightly to the left relative to the medial condition.

The findings for the σ parameter are presented in the middle panel of Figure 1. The Age × SW interaction was reliable, γ = 132; 95% HDI: [97, 168], but again, there was no evidence for an Age × CW interaction, γ = 24; 95% HDI: [−11, 59]. The Age × SW interaction was reliable because the difference in σ between the medial and sentence-final conditions was larger among older adults, γ = 212; 95% HDI: [181, 244], compared with the young, though the SW effect for the young adults was still different from zero, γ = 79; 95% HDI: [61, 97]. Collapsing across age groups, there was no reliable CW effect on σ, γ = 2; 95% HDI: [−16, 22].

For the τ parameter (Figure 1, Bottom panel), both the Age × SW interaction, γ = 63; 95% HDI: [29, 96], and Age × CW interaction, γ = 45; 95% HDI: [10, 77], were reliable. For younger adults, there was a reliable clause wrap-up effect on τ, with significant differences between the medial and clause positions, γ = 54; 95% HDI: [36, 71], as well as a sentence wrap-up effect, with a significant difference emerging between the medial and sentence final positions γ = 65; 95% HDI: [48, 82]. These effects were exaggerated among the older adults, who showed a larger difference in τ between the medial and clause-final positions, γ = 99; 95% HDI: [69, 129], and the medial and sentence-final positions, γ = 128; 95% HDI: [99, 159], compared with the young.

A graphical display of the modeled distributions is presented in Figure 2, which illustrates the density plots for the estimated ex-Gaussian parameters for each of the word position conditions for younger (left) and older adults (right). These are based on 40,000 random samples from each distribution, where each distribution is generated by summing a sample from a normal distribution with mean μ and standard deviation σ and a sample from an exponential distribution with rate parameter $1/\tau$. Sentence wrap-up appeared to shift the entire distribution, increase the variance in reading times, and increase the slow tail of the distribution (with greater effects for the older adults). However, clause wrap-up negatively shifted the RT distribution relative to medial words (i.e., median reading times were actually faster at clause boundaries) but, at the same time, there were selective positive effects on the τ parameter, suggesting that the presence of a marked clause boundary selectively increased RTs on the slowest trials, an effect that was larger for older adults.

Nonparametric Analyses

To assess distributional effects nonparametrically, vincentile plots, a descriptive method of examining RT distributions (Ratcliff, 1979), were constructed for each age group and each condition. Plots are constructed by rank ordering the residual reading times separately for each subject and condition and binning these residual RTs into deciles ranging from the fastest 10% of trials (vincentile 1) to the slowest 10% of trials (vincentile 10). Averages across all subjects are then plotted. This plot is presented in Figure 3. Effects that are attributable only to an overall shift in the distribution (e.g., an effect on μ) would result in a condition difference that is equal across all vincentiles (i.e., parallel lines). However, a shift in τ is seen in divergence at the slowest RTs (i.e., the largest vincentiles). As can be seen, the plot confirms the findings from the distributional fitting. For young adults, the sentence-final RTs are larger across all vincentiles, but diverge from sentence-medial RTs across larger vincentiles. However, clause-final RTs only diverge from the sentence-medial RTs at the largest vincentiles, corresponding to the right tail of the distribution. For older adults, there is a larger divergence from the sentence-medial condition for both the clause-final and sentence-final conditions at the slowest reading times.

Ex-Gaussian model estimated vincentiles are superimposed on the empirical vincentiles in Figure 3 (open triangles, circles, and

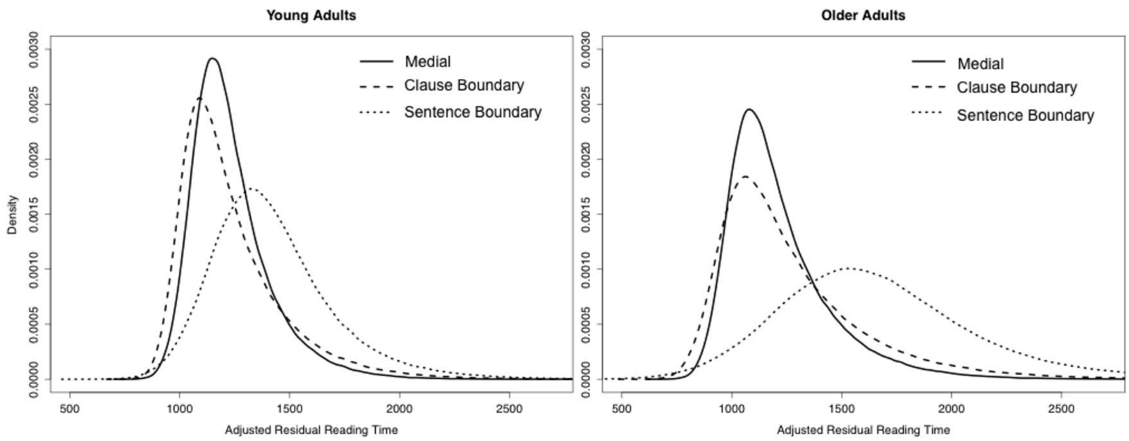


Figure 2. Simulated density functions from best fitting ex-Gaussian parameters for each word position for younger and older adults. Values are based on residual reading times that were rescaled by adding a constant to remove nonzero values for ex-Gaussian model fitting (see text for details).

squares). These were generated using the same process to generate the density plots in Figure 2. The degree to which the model estimated vincentile estimates approximate the empirical estimates indicates the degree of fit between the observed and model implied values. That these values nearly overlap over all points of the vincentiles for all conditions for both younger and older adults indicates good fit.

Individual Differences in Wrap-Up Distributions

Our final question addressed whether individual differences in cognitive and sensory ability were related to individual differences in the degree to which clause and sentence wrap-up impacted RT distribution components. Distributional influences of clause and sentence wrap-up were estimated by calculating separate clause wrap-up and sentence wrap-up effects for each parameter for each individual subject, using model estimated best-linear unbiased predictors (BLUPs) of the CW and SW contrasts[3] (Baayen et al., 2008). The resulting parameters represent change in each distributional parameter in the clause-final position (μ_{CW}, σ_{CW}, and τ_{CW}) and sentence-final position (μ_{SW}, σ_{SW}, and τ_{SW}) relative to the medial position (i.e., the clause wrap-up/sentence wrap-up effect for μ, σ, and τ).

The left hand side of Table 3 presents the intercorrelations among the clause wrap-up and sentence wrap-up parameters. For both clause and sentence wrap-up, the μ and σ parameters were positively correlated, suggesting that those who showed a larger shift in the distribution as a function of wrap-up also showed more variability in wrap-up. However, τ_{CW} was uncorrelated with μ_{CW} and σ_{CW}. A small negative correlation between μ_{SW} and τ_{SW} was found, suggesting that those who showed a larger shift in the distribution as a function of sentence wrap-up actually showed less of an effect of wrap-up on the slow tail of the distribution.

The right hand side of Table 3 contains correlations between the wrap-up distributional parameters and individual differences in age, visual acuity, working memory, speed of processing, and verbal ability. As seen in Table 3, sentence and clause wrap-up parameters were selectively correlated with individual differences in sensory and cognitive ability. For clause wrap-up, individuals with better vision showed less of a shift in the distribution, and less variability in RT distributions. However, verbal ability was positively correlated with distributional shifting and variance. A slower tail, on the other hand, was associated with poorer processing capacity, as indexed by measures of both speed and working memory. For sentence wrap-up, individuals with higher verbal ability, worse vision, and slower processing speed showed a larger shift and greater variability in RT distributions. A slower tail τ, on the other hand, was again associated with worse processing capacity, as indexed by measures of both speed and working memory.

In an attempt to disentangle these effects and examine whether individual differences in sensory and cognitive abilities might uniquely explain the observed age-related increases in μ, σ, and τ as a function of clause wrap-up and sentence wrap-up, we fit a multiple mediation model (see Hayes, 2013; Hayes & Scharkow, in press) with age treated as the focal predictor; visual acuity, verbal ability, processing speed, and working memory treated as

[3] Best linear unbiased predictors (BLUPs) are individual-specific estimators of item-level effects (e.g., wrap-up effects on reading-time distributions), estimated through model-derived parameters from a mixed effects model where all item-level predictors are allowed to vary randomly across subjects (i.e., where random slopes are estimated for all within-subject effects; see Baayen et al., 2008; and Snijders & Bosker, 2011 for an introduction to BLUP estimation in mixed models). BLUP estimates are shrunken toward the grand mean differentially more for extreme and less stable sample estimates, producing more reliable estimators of subject-specific effects than unadjusted effects (e.g., individual-level difference scores; see Gelman, Hill, & Yajima, 2012 for an explanation of the computational benefits of shrinkage in mixed models). Conceptually, the six parameters represent

$$\mu_{cw} = \mu_{clause} - \mu_{medical}$$
$$\mu_{sw} = \mu_{sentence} - \mu_{medical}$$
$$\sigma_{cw} = \sigma_{clause} - \sigma_{medical}$$
$$\sigma_{sw} = \sigma_{sentence} - \sigma_{medical}$$
$$\tau_{cw} = \tau_{clause} - \tau_{medical}$$
$$\tau_{sw} = \tau_{sentence} - \tau_{medical}$$

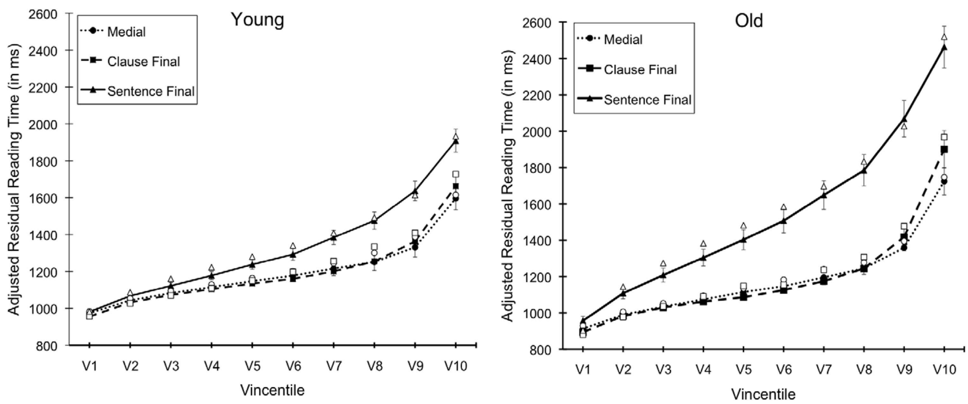


Figure 3. Empirical vincentile plots for each word position for younger and older adults. Model estimated values are overlaid (white fill) to assess model fit. Values are based on residual reading times that were rescaled by adding a constant to remove nonzero values for ex-Gaussian model fitting (see text for details).

mediators; and each of the six wrap-up distributional parameters treated as outcomes. Figure 4 presents a simplified path diagram of this multiple mediation model. We used the bootstrap procedure described in Preacher and Hayes (2008) to estimate indirect effects (resampling 5,000 times). These indirect effects represent tests for the degree to which each mediator variable uniquely explains the observed relationship between the focal predictor (age) and each outcome (the distributional parameters). Results of the multiple mediation model are shown in Table 4, which presents parameter estimates and 95% confidence intervals for each indirect effect.

Visual acuity was found to be a reliable mediator of both μ_{CW} and σ_{CW}, indicating that poorer visual acuity mediated the age-related influence of clause wrap-up on distributional shifting and scaling. Verbal ability had a significant, but negative indirect effect on τ_{CW}, indicating that age and verbal skill have opposing influences on the tail of the distribution (a case of inconsistent mediation,[4] see MacKinnon et al., 2000, 2007): aging was associated with increases in the slow tail of the distribution as a function of clause wrap-up, but greater verbal ability was associated with a reduced influence of clause wrap-up on the tail of the distribution.

Table 3
Correlations Among Ex-Gaussian BLUPs and Relationships With Individual Differences

Effect	1	2	3	4	5	Age	Vision	Verbal	Speed	WM
Clause										
1. μ_{CW}						**.15**	**−.16**	**.13**	−.07	−.12
2. σ_{CW}	**.87**					**.22**	**−.20**	**.18**	−.10	−.11
3. τ_{CW}	−.13	.04				**.31**	**−.17**	.10	**−.19**	**−.20**
Sentence										
4. μ_{SW}	**.32**	**.42**	**.43**			**.33**	**−.20**	**.36**	**−.15**	−.08
5. σ_{SW}	**.27**	**.33**	**.52**	**.82**		**.39**	**−.23**	**.35**	**−.20**	−.12
6. τ_{SW}	**−.17**	−.15	**.31**	**−.19**	.09	**.28**	−.10	.10	**−.30**	**−.21**

Note. Clause = Clause wrap-up effect; Sentence = Sentence wrap-up effect; $\mu_{CW/SW}$, $\sigma_{CW/SW}$, and $\tau_{CW/SW}$ estimates are individual BLUPs for difference between clause/sentence position and medial position (i.e., effects of wrap-up); WM = Working Memory. Bolded estimates are significant at $p < .05$.

Verbal ability was a reliable mediator of age differences in μ_{SW}, indicating that individuals with greater verbal knowledge showed a larger shift in the distribution at sentence boundaries. No individual differences were found to mediate the age-related increases in σ_{SW}. Verbal ability had a negative indirect effect on τ_{SW} (MacKinnon et al., 2000), indicating that individuals with better verbal ability showed a smaller proportion of extreme reading times at sentence boundaries. In contrast, both working memory and speed of processing were significant positive mediators of the effects of age on τ_{SW}, suggesting that age-related increases in τ at sentence boundaries can be explained in part by older adults' poorer working memory and processing speed.

Discussion

The current study investigated age and individual differences in distributions of reading times as a way to probe semantic integration mechanisms in reading. The ex-Gaussian model provided an excellent fit to self-paced reading time data for both younger and older adults, indicating that such parametric models provide useful descriptive summaries of RT distributions and distributional changes as a function of both item-level (e.g., boundary position) and subject-level (e.g., age, working memory) variability. In the following sections, we consider (a) how RT distributional analyses provide insights into models of wrap-up in sentence processing, (b) aging and individual differences in sensory and cognitive function and RT distributions, and (c) limitations of the current study and future directions.

[4] MacKinnon et al. (2007; MacKinnon, Krull & Lockwood, 2000) have used the term *inconsistent mediation* to describe cases in tests of mediation models where the direct and mediated effects of an independent variable on a dependent variable have opposite signs (see also Davis, 1985; McFatter, 1979). Such a case exists for the effects of verbal ability on τ_{CW} and τ_{SW}. Aging has a positive path (a) to verbal ability, and a negative path (b) to $\tau_{CW/SW}$, generating a negative estimate of the indirect effect (ab). This is similar to suppression effects in multiple regression models (see Mackinnon et al., 2000) and indicates that the proposed focal predictor (age) and the mediator (verbal ability) have independent and opposing influences on the outcome ($\tau_{CW/SW}$).

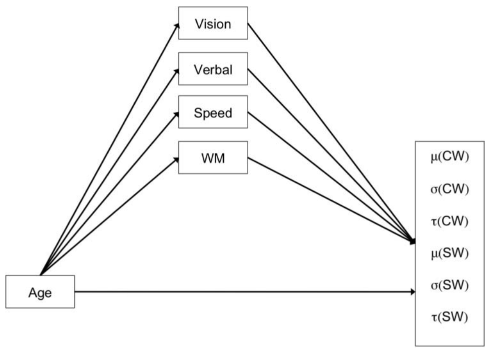


Figure 4. Simplified path diagram of multiple mediator model of age and individual differences in ex-Gaussian parameters for clause and sentence wrap-up.

Ex-Gaussian Distributions: Implications for Models of Wrap-Up

A major goal of the current study was to utilize RT distributional methods to examine how sentence comprehension processes (e.g., wrap-up) may influence underlying distributions of reading times for different individuals. We were motivated to use this approach to distinguish between automatic/obligatory models of wrap-up and semantic integration models of wrap-up because, although both classes of models predict increases in mean RT at clause and sentence boundaries, these theories imply different distributional properties.

Automatic/obligatory models of wrap-up predict that pause durations at clause-boundaries should occur equally across short and long reading times, consistent with the idea that boundary sites trigger an automatic "dwell" response that is not dependent upon properties of the text (Hirotani et al., 2006; Warren et al., 2009). In contrast, the findings from the current study showed that wrap-up not only engendered a uniform increase in reading time, but also increased variability in reading times at sentence boundaries and an increased proportion of extreme reading times at both clause and sentence boundaries. The exaggerated variability and slow tail in reading time distributions for words at syntactic boundaries suggests that the effects of wrap-up are not fixed or obligatory. These findings are not easily compatible with theories of wrap-up that explain increased pause durations at boundary sites in terms of automatic mechanisms, such as an oculomotor pause response or the automatic monitoring of intonational boundaries alone.

To our knowledge, this is the first study examining the influence of sentence comprehension processes on RT distributions in self-paced reading (but see Staub, 2011, for an example in eye tracking). Note that, although the above findings cast doubt on the idea that wrap-up effects are uniform in nature, they do not unequivocally implicate a cognitively effortful integration process. In the next section, we consider how our aging and individual differences findings suggest that wrap-up effects are cognitively effortful.

Age and Individual Differences in Distributions of Wrap-Up

Older adults showed exaggerated shifting, scaling, and slow-tail effects for sentence final RTs, as well as larger slow-tail effects at clause boundaries relative to nonboundary sites. Previously, we (Payne & Stine-Morrow, 2012) found that age differences in wrap-up in eye-tracking were largest for trials that included both refixations and regressions (see also Rayner et al., 2009, 2010) and suggested that the difference in findings between early pass (e.g., first fixation duration) and late-pass (e.g., regression path duration) fixation measures could be attributable to a subset of trials where wrap-up was highly demanding, and thus where more time was needed to process the target word. A major advantage of the ex-Gaussian model is the ability to examine condition differences across the entire RT distribution, rather than relying on inferring such effects by comparing mean RTs across several outcome measures (e.g., Payne & Stine-Morrow, 2012; Rayner et al., 2009, 2010). The findings in the current study were consistent with the claim that age differences in sentence wrap-up were exaggerated on the most demanding words. That is, age differences in reading time were not uniform in nature, but varied as a function of item difficulty.

Results from our individual difference analyses also suggested that wrap-up effects are cognitively effortful. Processing speed, working memory, and verbal ability differentially influenced both clause wrap-up and sentence wrap-up effects on reading time distributions. Most notably, working memory and processing speed were more highly correlated with τ effects of clause and sentence wrap-up than with other RT parameters (see Table 3). Our mediational analyses showed that individual differences in cognitive ability independently mediated age differences apart

Table 4
Parameter Estimates and 95% Confidence Intervals of Direct and Indirect Effects From Multiple Mediation Model

	Clause wrap-up			Sentence wrap-up		
Effect	μ_{CW}	σ_{CW}	τ_{CW}	μ_{SW}	σ_{SW}	τ_{SW}
DE						
Age	−.07 [−.55, .42]	.37 [−.21, .95]	**2.01 [1.22, 2.91]**	2.13 [−1.14, 5.40]	**2.67 [.82, 2.92]**	**2.02 [.67, 3.37]**
IE						
Vision	**.09 [.01, .21]**	**.13 [.03, .25]**	.04 [−.19, .25]	.31 [−.55, 1.18]	.13 [−.38, .62]	−.24 [−.64, .14]
WM	.09 [−.04, .23]	−.01 [−.19, .16]	−.16 [−.45, .12]	−.80 [−2.01, .22]	−.59 [−1.29, .01]	**.18 [.08, .31]**
Speed	−.001 [−.16, .15]	−.06 [−.24, .09]	−.12 [−.42, .17]	−.06 [−1.19, 1.09]	−.001 [−.67, .66]	**.43 [.01, .88]**
Verbal	.17 [−.11, .65]	.04 [−.27, .60]	**−.96 [−1.52, −.44]**	**2.31 [.21, 4.50]**	.39 [−.75, 1.67]	**−.90 [−1.96, −.03]**

Note. DE = direct effect; IE = indirect effect; WM = working memory; $\mu_{CW/SW}$, $\sigma_{CW/SW}$, and $\tau_{CW/SW}$ estimates are individual BLUPs for difference between clause/sentence position and medial position (i.e., effects of wrap-up); Values in brackets are the parametric 95% CI for direct effects and the 95% bias corrected bootstrapped CIs for indirect effects. Bolded estimates are significant at $p < .05$.

from effects of sensory ability. In particular, our findings indicated that age-related changes in mean RT at clause and sentence boundaries were driven by a number of processes that work in opposing directions and on different components of the RT distribution. Older adults with greater verbal ability showed larger shifting in the RT distribution as a function of wrap-up, along with a smaller proportion of extreme reading times at boundary sites. In contrast, age-related reductions in working memory and processing speed were responsible in part for the increase in extreme reading times at clause boundaries and sentence boundaries.

Overall, our age-related and individual difference findings suggested that the more pronounced tails at clause and sentence boundaries reflect a cognitively demanding process that becomes more frequently disrupted with advancing age and with poorer working memory and psychomotor speed. These results are consistent with Staub and Benatar's (2013) explanation for dissociations between μ and τ in eye-movement studies of sentence reading. They claimed that distributional shifting may result in local increases in processing difficulty, whereas increases in proportions of extreme RTs likely reflect an increased proportion of trials where normal processing was disrupted, consistent with claims made in the attentional control literature (Coyle, 2003; Tse et al., 2010). Under this account, slower tails at boundary sites implicate increased rates of disruption in wrap-up with advancing age and poorer fluid cognitive function.

The degree to which working memory impacts online language processing is highly contested (Caplan & Waters, 1999; Caplan, Waters, & DeDe, 2007; MacDonald & Christiansen, 2002; Just & Varma, 2007). Our findings suggest that individual differences in WM may have varying degrees of influence on particular RT components in online language processing, with stronger effects for the reading times in the tail of the distribution. To the extent that certain linguistic variables impact μ, σ, and τ, this may help to explain the observed variability in detection of working memory effects across studies of online sentence processing. Individual differences in working memory may be more likely to be detected for linguistic manipulations that increase the slow tail of a distribution, rather than simply shifting the distribution. Further research examining the influences of other aspects of linguistic difficulty on individual differences in RT distributions is necessary to validate this claim.

Limitations and Future Directions

Important limitations of the current study need to be addressed. We also explore avenues for future work on aging and RT distributions and applications of RT distribution analysis to sentence processing.

First, this study adopted a corpus-based approach to examine individual differences in the effects of wrap-up on sentence processing, using naturalistic sentences. Corpus-based approaches have been recently gaining in popularity and have provided valuable data on naturalistic sentence processing (Kuperman & Van Dyke, 2011; Kuperman, Dambacher, Nuthmann & Kliegl, 2010; Kliegl, Grabner, Rolfs, & Engbert, 2004; Nuthmann & Kliegl, 2009) and visual word recognition (Balota et al., 2007; Dufau et al., 2011; Yap et al., 2012). However, such findings are necessarily correlational. Despite our efforts to control for potentially confounding item-level influences statistically (see also Kuperman et al., 2010), we cannot be sure that the findings in the current study were completely independent of other item-level linguistic influences that covary with relative word position. In future research, distributional effects of wrap-up should be examined experimentally, by manipulating clause and sentence boundaries across matched sentence material, to completely control for other potential linguistic influences (e.g., Payne & Stine-Morrow, 2012; Rayner et al., 2000). Despite this constraint, the current findings are valuable in allowing us to not only examine item-level distributional influences, but also individual differences in RT distributions with a large and diverse sample that is adequately powered to detect individual differences in RT distributions.

Second, unlike in simple two-choice decision paradigms (Ratcliff & McKoon, 2008), there is currently no computationally explicit process model to link reading time distributions to underlying psychological phenomena in sentence processing. However, this does not preclude the use of RT distributional analyses to investigate reading and fixation time data beyond the influence of variables on the mean (see Balota & Yap, 2011 for a discussion). Such descriptive methods allow for a fine-grained examination of the effects of both item-level manipulations and individual differences on RTs (see Balota et al., 2008; Balota & Yap, 2011). Another advantage of distributional analyses is that, as a statistical method, they can be applied across a broader range of tasks and paradigms. Indeed, the current findings provide initial empirical data on the influence of sentence comprehension processes on underlying RT distributions in self-paced reading, which is critical for the development of fine grained computational models that not only predict condition differences in mean RT, but can account for distributional influences.

Current computational models in sentence processing focus on mean reading time across conditions, without consideration of influences on underlying RT distribution shape (MacDonald & Christiansen, 2002; Just & Varma, 2002; Pollatsek et al., 2006; Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; Gibson, 1998). Such critiques of this "aggregating" approach have been previously voiced (Feng, 2003). This aggregation approach in modeling is likely driven by a lag in empirical research reporting effects of linguistic influences on RT distributions in self-paced reading and eye-movement studies. To the extent that researchers continue to examine effects of linguistic difficulty on reading and fixation time distributions (e.g., Staub, 2011), such findings will likely lead to more refined computational models, capable of predicting not only mean RTs but underlying distributions as well.

Lastly, an important goal for future work is to examine the degree to which RT distributions differ across paradigms, including self-paced reading and eye-movement experiments. Some have argued that choices in text presentation and artificial motor components in the self-paced reading paradigm result in distorted effects (Adams, Clifton, & Mitchell, 1998; Rayner, 1998; Magliano et al., 1993). However, Mitchell (2004) argued that "there is no solid evidence that researchers have been misled by segmentation biases in the self-paced reading task" (p. 26). Indeed, a number of studies have replicated influences of linguistic complexity across self-paced reading and eye tracking at the mean level, including relative clause attachment (Carreiras & Clifton, 1993; Traxler et al., 1998), word frequency (Mitchell & Green, 1978; Rayner, 1977) and wrap-up effects (Aaronson & Scarborough, 1976; Rayner et al., 2000). It has been argued that wrap-up effects in eye tracking are smaller than in self-paced reading in at least one study (Magliano et al., 1993). However, other researchers have found robust effects of wrap-up in eye tracking experiments with younger and older adults (Hirotani et al., 2006; Payne et al., 2012; Rayner et al., 2000; Stine-Morrow et al., 2010).[5]

Conclusion

The current study is, to our knowledge, the first to apply a distributional analysis to examine self-paced word-by-word sentence reading data. Overall, we found that both parametric and nonparametric distributional analytical tools provide insight into the effects of aging and individual differences in cognition and sensory ability on online sentence comprehension, beyond traditional measures that rely primarily on mean RT. In particular, our findings were consistent with integration views of wrap-up and suggested that age and individual differences in attentional allocation during reading, as revealed by RT distribution analyses, play an important role in sentence understanding.

[5] This is particularly important to consider in studies comparing age differences in language processing. Older adults have been argued to be more "risky" readers (Rayner et al., 2006), showing larger skipping rates and more regressions during reading, effects that are obscured in the moving window paradigm. Stine-Morrow and colleagues (2010) directly compared age differences in wrap-up across eye tracking and self-paced reading. They showed that age differences in the effects of wrap-up were comparable across both self-paced reading and eye tracking with the same stimuli. Thus, at the mean level, age differences in wrap-up have been shown to be equivalent across methodologies. However, future work should investigate whether this equivalence is also observed at the level of RT distributions.

References

Aaronson, D., & Ferres, S. (1984). The word-by-word reading paradigm: An experimental and theoretical approach. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 31–68). Hillsdale, NJ: Erlbaum.

Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: Some qualitative evidence. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 56–70. doi:10.1037/0096-1523.2.1.56

Adams, B. A., Clifton, C., Jr., & Mitchell, D. C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin & Review, 5,* 265–270. doi:10.3758/BF03212949

Anstey, K. J., Hofer, S. M., & Luszcz, M. A. (2003). A latent growth curve analysis of late-life sensory and cognitive function over 8 years: Evidence for specific and common factors underlying change. *Psychology and Aging, 18,* 714–726. doi:10.1037/0882-7974.18.4.714

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. doi:10.1016/j.jml.2007.12.005

Balogh, J., Zurif, E., Prather, P., Swinney, D., & Finkel, L. (1998). Gap-filling and end of sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language, 61,* 169–182. doi:10.1006/brln.1997.1917

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128,* 32–55. doi:10.1037/0096-3445.128.1.32

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20,* 160–166. doi:10.1177/0963721411408885

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459. doi:10.3758/BF03193014

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59,* 495–523. doi:10.1016/j.jml.2007.10.004

Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes. http://CRAN.R-project.org/package=lme4

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 373–443). New York, NY: Psychology Press.

Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception & Psychophysics, 12,* 73–76. doi:10.3758/BF03212846

Caplan, D., DeDe, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging, 26,* 439–450. doi:10.1037/a0021837

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77–94. doi:10.1017/S0140525X99001788

Caplan, D., & Waters, G. S. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin & Review, 20,* 243–268. doi:10.3758/s13423-012-0369-9

Caplan, D., Waters, G., & DeDe, G. (2007). Specialized verbal working memory for language comprehension. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory*. New York, NY: Oxford University Press.

Carreiras, M., & Clifton, C., Jr. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech, 36,* 353–372.

Castel, A. D., Balota, D. A., Hutchison, K. A., Logan, J. M., & Yap, M. J. (2007). Spatial attention and response control in healthy younger and older adults and individuals with Alzheimer's disease: Evidence for disproportionate selection breakdowns in the Simon task. *Neuropsychology, 21,* 170–182. doi:10.1037/0894-4105.21.2.170

Cheung, H., & Kemper, S. (1992). Competing complexity metrics and adults' production of complex sentences. *Applied Psycholinguistics, 13,* 53–76. doi:10.1017/S0142716400005427

Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. *Intelligence, 31,* 567–587. doi:10.1016/S0160-2896(03)00054-0

Criss, A. H. (2010). Differentiation and response bias in episodic memory: Evidence from reaction time distributions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 484–499. doi:10.1037/a0018435

Dagerman, K. S., MacDonald, M. C., & Harm, M. W. (2006). Aging and the use of context in ambiguity resolution: Complex changes from simple slowing. *Cognitive Science, 30,* 311–345. doi:10.1207/s15516709cog0000_46

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 561–584. doi:10.1037/0278-7393.9.4.561

Davis, M. D. (1985). The logic of causal order. In J. L. Sullivan and R. G. Niemi (Eds.), *Sage university paper series on quantitative applications in the social sciences*. Beverly Hills, CA: Sage.

DeDe, G., Caplan, D., Kemtes, K., & Waters, G. (2004). The relationship between age, verbal working memory and language comprehension. *Psychology and Aging, 19,* 601–616. doi:10.1037/0882-7974.19.4.601

Dufau, S., Duñabeitia, J. A., Moret-Tatay, C., McGonigal, A., Peeters, D., Alario, F.-X., . . . Grainger, J. (2011). Smart phone, smart science: How the use of smartphones can revolutionize research in cognitive science. *PLoS ONE, 6,* e24974. doi:10.1371/journal.pone.0024974

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for the kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Fallon, M., Peelle, J. E., & Wingfield, A. (2006). Spoken sentence processing in young and older adults modulated by task demands: Evidence from self-paced listening. *Journal of Gerontology: Psychological Sciences, 61*B, 10–17. doi:10.1093/geronb/61.1.P10

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44,* 491–505. doi:10.1111/j.1469-8986.2007.00531.x

Feng, G. (2003). Throwing the baby out with the bathwater: Problems in modeling aggregated eye-movement data. *Behavioral and Brain Sciences, 26,* 482–483. doi:10.1017/S0140525X03280109

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348–368. doi:10.1016/0749-596X(86)90006-9

Ferreira, F., & Patson, N. (2007). The good enough approach to language comprehension. *Language and Linguistics Compass, 1,* 71–83. doi:10.1111/j.1749-818X.2007.00007.x

Fodor, J., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning & Verbal Behavior, 4,* 414–420. doi:10.1016/S0022-5371(65)80081-0

Garrett, M., Bever, T. G., & Fodor, J. (1966). The active use of grammar in speech perception. *Perception & Psychophysics, 1,* 30–32. doi:10.3758/BF03207817

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness, 5,* 189–211. doi:10.1080/19345747.2011.618213

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68,* 1–76. doi:10.1016/S0010-0277(98)00034-1

Haberlandt, K., & Graesser, A. C. (1989). Processing new arguments at clause boundaries. *Memory & Cognition, 17,* 186–193. doi:10.3758/BF03197068

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15,* 883–899. doi:10.1162/089892903322370807

Hayes, A. (2013). *Introduction to mediation, moderation, and conditional process analysis*. New York, NY: The Guilford Press.

Hayes, A. F. & Scharkow, M. (in press). The relative trustworthiness of tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science*.

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers, 36,* 277–290. doi:10.3758/BF03195574

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9,* 394–401. doi:10.3758/BF03196299

Henderson, J. M., & Ferreria, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429. doi:10.1037/0278-7393.16.3.417

Hill, R., & Murray, W. (2000). Commas and spaces: Effects of punctuation on eye movements and sentences processing. In A. Kennedy, D. Heller, & J. Pynte (Eds.) *Reading as a perceptual process* (pp. 565–589). Amsterdam, The Netherlands: Elsevier. doi:10.1016/B978-008043642-5/50027-9

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54,* 425–443. doi:10.1016/j.jml.2005.12.001

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes, 26,* 376–405. doi:10.1080/01690965.2010.492642

Hultsch, D. F., Strauss, E., Hunter, M. A., & MacDonald, S. (2008). Intraindividual variability, cognition, and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 491–556). New York, NY: Psychology Press.

Jackson, J. D., Balota, D. A., Duchek, J. M., & Head, D. (2012). White matter integrity and reaction time intraindividual variability in healthy aging and early-stage Alzheimer disease. *Neuropsychologia, 50,* 357–366. doi:10.1016/j.neuropsychologia.2011.11.024

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61,* 23–62. doi:10.1016/j.cogpsych.2010.02.002

Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning & Verbal Behavior, 10,* 409–416. doi:10.1016/S0022-5371(71)80040-3

Johnson, R. E. (2003). Aging and the remembering of text. *Developmental Review, 23,* 261–346. doi:10.1016/S0273-2297(03)00009-1

Johnson, R. L., Staub, A., & Fleri, A. M. (2012). Distributional analysis of the transposed-letter neighborhood effect on naming latency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1773–1779. doi:10.1037/a0028222

Just, M. A., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354. doi:10.1037/0033-295X.87.4.329

Just, M. A., & Varma, S. (2002). A hybrid architecture for working memory: Reply to MacDonald and Christiansen (2002). *Psychological Review, 109,* 55–65. doi:10.1037/0033-295X.109.1.55

Just, M. A., & Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. *Cognitive, Affective, & Behavioral Neuroscience, 7,* 153–191. doi:10.3758/CABN.7.3.153

Kemper, S., & Liu, C. (2007). Eye movements of young and older adults during reading. *Psychology and Aging, 22,* 84–93. doi:10.1037/0882-7974.22.1.84

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition, 2,* 15–47. doi:10.1016/0010-0277(72)90028-5

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16,* 262–284. doi:10.1080/09541440340000213

Kruschke, J. K. (2011). *Doing Bayesian data analysis: A tutorial with R and BUGS*. San Diego, CA: Academic Press/Elsevier.

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *The Quarterly Journal of Experimental Psychology, 63,* 1838–1857. doi:10.1080/17470211003602412

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65,* 42–73. doi:10.1016/j.jml.2011.03.002

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29,* 375–419. doi:10.1207/s15516709cog0000_25

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10,* 447–454. doi:10.1016/j.tics.2006.08.007

Lien, M. C., Allen, P. A., Ruthruff, E., Grabbe, J., McCann, R. S., & Remington, R. W. (2006). Visual word recognition without central attention: Evidence for greater automaticity with advancing age. *Psychology and Aging, 21,* 431. doi:10.1037/0882-7974.21.3.431

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9,* 339–355. doi:10.1037/0882-7974.9.3.339

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

Luo, Y., Yan, M., & Zhou, X. (2013). Prosodic boundaries delay the processing of upcoming lexical information during silent sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 915. doi:10.1037/a0029182

MacDonald, M. C., & Christiansen, M. C. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109,* 35–54. doi:10.1037/0033-295X.109.1.35

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58,* 593. doi:10.1146/annurev.psych.58.110405.085542

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1,* 173–181. doi:10.1023/A:1026595011371

Magliano, J. P., Graesser, A. C., Eymard, L. A., Haberlandt, K., & Ghoulson, B. (1993). Locus of interpretive and inference processes during text comprehension: A comparison and gaze durations and word reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 704–709. doi:10.1037/0278-7393.19.3.704

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of ex–Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16,* 798–817. doi:10.3758/PBR.16.5.798

McFatter, R. M. (1979). The use of structural equation models in interpreting regression equations including suppressor and enhancer variables. *Applied Psychological Measurement, 3,* 123–135. doi:10.1177/014662167900300113

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to "D'oh!" Working memory capacity and mind wandering predict extreme reaction times and executive-control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 525–549. doi:10.1037/a0025896

Miller, L. M. S., & Stine-Morrow, E. A. L. (1998). Aging and the effects of knowledge on online reading strategies. *Journal of Gerontology: Psychological Sciences, 53B,* P223–P233. doi:10.1093/geronb/53B.4.P223

Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language, 33,* 128–147. doi:10.1006/jmla.1994.1007

Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond* (pp. 15–32). New York, NY: Psychology Press.

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content of immediate processing in reading. *Quarterly Journal of Experimental Psychology, 30,* 609–636. doi:10.1080/14640747808400689

Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision, 9,* 31. doi:10.1167/9.5.31

Payne, B. R., Gao, X., Noh, S. R., Anderson, C. J., & Stine-Morrow, E. A. L. (2012). The effects of print exposure on sentence processing and memory among older adults: Evidence for efficiency and reserve. *Aging, Neuropsychology, and Cognition, 19,* 122–149. doi:10.1080/13825585.2011.628376

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. L. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition, 130,* 157–173. doi:10.1016/j.cognition.2013.10.005

Payne, B. R., Gross, A., Parisi, J., Sisco, S., Stine-Morrow, E. A. L., Marsiske, M., & Rebok, G. (2014). Modeling longitudinal changes in older adults' memory for spoken discourse: Findings from the ACTIVE cohort. *Memory*.

Payne, B. R., & Stine-Morrow, E. A. L. (2012). Aging, parafoveal preview, and semantic integration in sentence processing: Testing the cognitive workload of wrap-up. *Psychology and Aging, 27,* 638–649. doi:10.1037/a0026540

Pollatsek, A., Reichle, E. D.,& Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52,* 1–56. doi:10.1016/j.cogpsych.2005.06.001

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879–891. doi:10.3758/BRM.40.3.879

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication, 43,* 103–121. doi:10.1016/j.specom.2004.02.004

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461. doi:10.1037/0033-2909.86.3.446

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922. doi:10.1162/neco.2008.12-06-420

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition, 5,* 443–448. doi:10.3758/BF03197383

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422. doi:10.1037/0033-2909.124.3.372

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging, 24,* 755–760. doi:10.1037/a0014300

Rayner, K., Castelhano, M. S., & Yang, J. (2010). Preview benefit during eye fixations in reading for older and younger readers. *Psychology and Aging, 25,* 714–718. doi:10.1037/a0019199

Rayner, K., & Clifton, C. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology, 80,* 4–9. doi:10.1016/j.biopsycho.2008.05.002

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53,* 1061–1080. doi:10.1080/713755934

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

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

Rohrer, D., & Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition, 22,* 511–524. doi:10.3758/BF03198390

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27*(5), 763. doi:10.1037/0012-1649.27.5.763

Salthouse, T. A., & Berish, D. E. (2005). Correlates of within-person (across-occasion) variability in reaction time. *Neuropsychology, 19,* 77–87. doi:10.1037/0894-4105.19.1.77

Schmiedek, F., Oberauer, K., Wilhelm, O., Süss, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136,* 414–429. doi:10.1037/0096-3445.136.3.414

Schroeder, S. (2011). What readers have and do: Effects of students' verbal ability and reading time components on comprehension with and without text availability. *Journal of Educational Psychology, 103,* 877–896. doi:10.1037/a0023731

Sharkey, N. E., & Sharkey, A. J. C. (1987). What is the point of integration? The loci of knowledge-based facilitation in sentence processing. *Journal of Memory and Language, 26,* 255–276. doi:10.1016/0749-596X(87)90114-8

Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford.

Smiler, A. P., Gagne, D. D., & Stine-Morrow, E. A. L. (2003). Aging, memory load, and resource allocation during reading. *Psychology and Aging, 18,* 203–209. doi:10.1037/0882-7974.18.2.203

Snijders, T. A. B., & Bosker, R. J. (2011). *Multilevel analysis*. London, UK: Sage.

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of reaction time distributions. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 506–526. doi:10.1037/0096-1523.26.2.506

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging, 25,* 377–390. doi:10.1037/a0018022

Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review, 18,* 371–376. doi:10.3758/s13423-010-0046-9

Staub, A., & Benatar, A. (2013). Individual differences in fixation duration distributions in reading. *Psychonomic Bulletin & Review, 20,* 1304–1311. doi:10.3758/s13423-013-0444-x

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1280–1293. doi:10.1037/a0016896

Stine, E. A. L. (1990). On-line processing of written text by younger and older adults. *Psychology and Aging, 5,* 68–78. doi:10.1037/0882-7974.5.1.68

Stine, E. A. L., & Hindman, J. (1994). Age differences in reading time allocation for propositionally dense sentences. *Aging, Neuropsychology, and Cognition, 1,* 2–16.

Stine-Morrow, E. A., Loveless, M. K., & Soederberg, L. M. (1996). Resource allocation in on-line reading by younger and older adults. *Psychology and Aging, 11,* 475. doi:10.1037/0882-7974.11.3.475

Stine-Morrow, E. A. L., & Miller, L. M. S. (2009). Aging, self-regulation, and learning from text. *Psychology of Learning and Motivation, 51,* 255–296. doi:10.1016/S0079-7421(09)51008-0

Stine-Morrow, E. A. L., Miller, L. M. S., Gagne, D. D., & Hertzog, C. (2008). Self- regulated reading in adulthood. *Psychology and Aging, 23,* 131–153. doi:10.1037/0882-7974.23.1.131

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

Stine-Morrow, E. A. L., Millinder, L., Pullara, P., & Herman, B. (2001). Patterns of resource allocation are reliable among younger and older readers. *Psychology and Aging, 16,* 69–84. doi:10.1037/0882-7974.16.1.69

Stine-Morrow, E. A. L., Shake, M. C., Miles, J. R., Lee, K., Gao, X., & McConckie, G. (2010). Pay now or pay later: Aging and the role of boundary salience in self-regulation of conceptual integration in sentence processing. *Psychology and Aging, 25,* 168–176. doi:10.1037/a0018127

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

Traxler, M. J. (2009). A hierarchical linear modeling analysis of working memory and implicit prosody in the resolution of adjunct attachment ambiguity. *Journal of Psycholinguistic Research, 38,* 491–509. doi:10.1007/s10936-009-9102-x

Traxler, M., Pickering, M., & Clifton, C., Jr. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39,* 558–592. doi:10.1006/jmla.1998.2600

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33,* 285–318. doi:10.1006/jmla.1994.1014

Tse, C. S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology, 24,* 300–315. doi:10.1037/a0018274

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7,* 424–465. doi:10.3758/BF03214357

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging, 18,* 332. doi:10.1037/0882-7974.18.2.332

Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E-Z reader. *Cognition, 111,* 132–137. doi:10.1016/j.cognition.2008.12.011

West, R. (2001). The transient nature of executive control processes in younger and older adults. *European Journal of Cognitive Psychology, 13,* 91–105. doi:10.1080/09541440042000232

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review, 12,* 891–896. doi:10.3758/BF03196782

White, S. J., & Staub, A. (2012). The distribution of fixation durations during reading: Effects of stimulus quality. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 603–617. doi:10.1037/a0025338

White, S. J., Warren, T., & Reichle, E. D. (2011). Parafoveal preview during reading: Effects of sentence position. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1221. doi:10.1037/a0022190

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. doi:10.3758/BF03209349

Yamani, Y., Chin, J., Meyers, E. A. G., Gao, X., Morrow, D. G., Stine-Morrow, E. A., . . . Murray, M. D. (2012). Reading engagement offsets declines in processing capacity for health literacy. *Proceedings of the 56th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 916–920). Boston, MA: Human Factors and Ergonomics Society.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 274–296. doi:10.1037/0278-7393.33.2.274

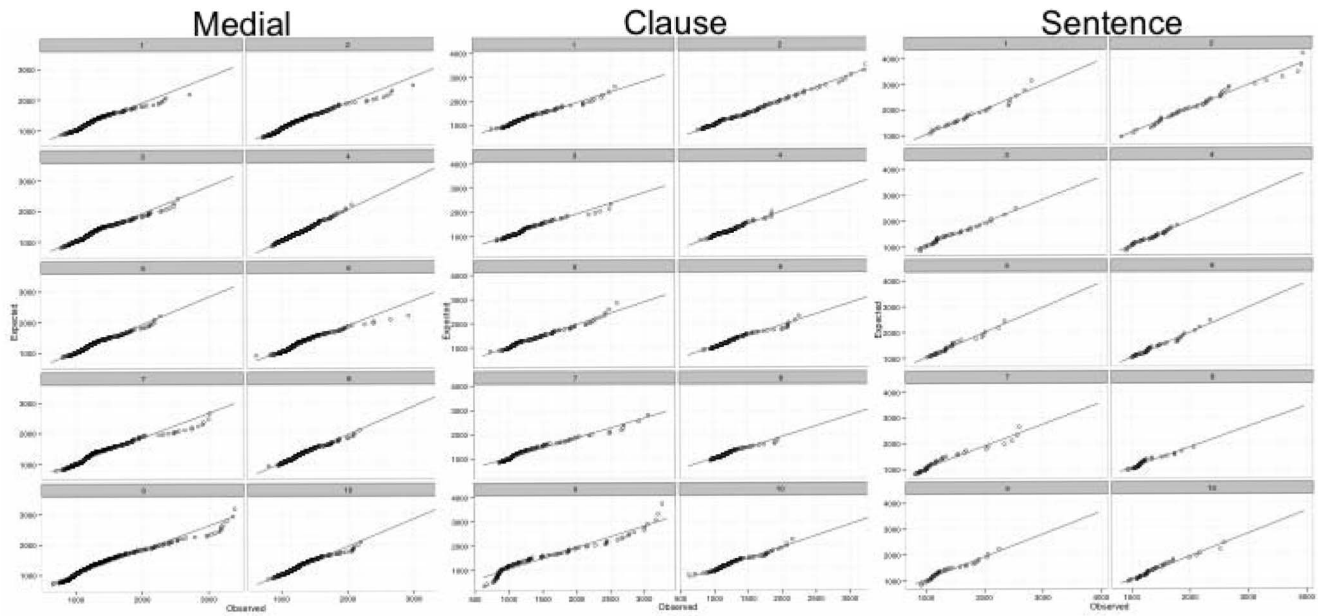
Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 53–79. doi:10.1037/a0024177

(*Appendix follows*)

Appendix

Graphic Display of Model Fit

Ex-Gaussian model expected RTs by observed RTs for randomly selected participants (n = 10) for medial, clause final, and sentence final words.



Received March 12, 2013
Revision received September 9, 2013
Accepted January 15, 2014 ■