The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

I "hear" what you're "saying": Auditory perceptual simulation, reading speed, and reading comprehension

Peiyun Zhou[ab] & Kiel Christianson[ab]
[a] Department of Educational Psychology, University of Illinois, Urbana-Champaign, Urbana, IL, USA
[b] Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, Urbana, IL, USA
Accepted author version posted online: 13 Feb 2015.Published online: 17 Mar 2015.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2015
http://dx.doi.org/10.1080/17470218.2015.1018282

Routledge Taylor & Francis Group

I "hear" what you're "saying": Auditory perceptual simulation, reading speed, and reading comprehension

Peiyun Zhou[1,2] **and Kiel Christianson**[1,2]

[1]Department of Educational Psychology, University of Illinois, Urbana-Champaign, Urbana, IL, USA
[2]Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, Urbana, IL, USA

(Received 8 August 2014; accepted 9 February 2015)

Auditory perceptual simulation (APS) during silent reading refers to situations in which the reader actively simulates the voice of a character or other person depicted in a text. In three eye-tracking experiments, APS effects were investigated as people read utterances attributed to a native English speaker, a non-native English speaker, or no speaker at all. APS effects were measured via online eye movements and offline comprehension probes. Results demonstrated that inducing APS during silent reading resulted in observable differences in reading speed when readers simulated the speech of faster compared to slower speakers and compared to silent reading without APS. Social attitude survey results indicated that readers' attitudes towards the native and non-native speech did not consistently influence APS-related effects. APS of both native speech and non-native speech increased reading speed, facilitated deeper, less good-enough sentence processing, and improved comprehension compared to normal silent reading.

Keywords: Auditory perceptual simulation; Embodied cognition; Good-enough processing; Language comprehension; Reading; Eye movements.

Syntactic structure is fragile, susceptible to both decay in memory (Sachs, 1967) and interference from competing lexical–semantic information (Christianson, Luke, & Ferreira, 2010; Ferreira, 2003; Lim & Christianson, 2013a, 2013b). Consequently, misinterpretations often arise when noncanonical syntactic structure is used to convey semantically implausible information. For example, readers regularly misinterpret both passive sentences (Example 1, below) and object-relative clauses (Example 2, below) to mean "the bird ate the worm". Ferreira (2003) proposes that such misinterpretations derive from a word-order heuristic in English that interprets, based on frequency and/or syntactic canonicity, noun–verb–noun strings as subject–verb–object structures with agent–verb–patient thematic alignments (Townsend & Bever, 2001). The interpretation derived from the application of this heuristic is then supported by general world knowledge, which, in the case of Examples 1–2 below, tells us that birds usually eat worms, rather than the other way around. Ferreira does not argue that the syntactic structure is not computed. Rather,

Correspondence should be addressed to Peiyun Zhou, Department of Educational Psychology, Education Building, MC210, University of Illinois, 1310 S. Sixth Street, Champaign, IL 61820, USA. E-mail: pzhou5@illinois.edu

We are grateful for the excellent work of the members of the Educational Psychology Psycholinguistics Lab, Beckman Institute for Advanced Science and Technology, who assisted with data collection, especially Cassie Palmer, Heather Mauch, Nayoung Kim, and Mallory Stites. We also thank Ben Swets and an anonymous reviewer for their feedback.

This work was supported by the National Science Foundation (NSF) [CAREER Award BCS-0847533 to Kiel Christianson].

the syntax is fragile and, under certain circumstances, is overridden by competing plausibility and heuristic-based processing. Substantial related research on visual language processing (i.e., reading) shows that readers often misinterpret sentences like Examples 1–2 as well as other difficult structures (such as so-called garden-path sentences, Example 3), despite signs of both rereading and coexisting partially correct interpretations (Christianson, 2008; Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Christianson & Luke, 2011; Christianson, Williams, Zacks, & Ferreira, 2006; Lim & Christianson, 2013a, 2013b; Patson, Darowski, Moon, & Ferreira, 2009; Slattery, Sturt, Christianson, Yoshida, & Ferreira, 2013; Swets, Desmet, Clifton, & Ferreira, 2008; van Gompel, Pickering, Pearson, & Jacob, 2006).

1. The bird was eaten by the worm.
2. The bird that the worm ate was slow.
3. While Anna dressed the baby that was cute played in the crib.

This line of research, all of which falls within the "good-enough processing" framework (Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007), has so far focused on documenting misinterpretations and accounting for how they arise. Importantly, the evidence suggests that these types of misinterpretations are not haphazard and do not stem from simple inattention. Instead, misinterpretations are predictable and systematic and derive from competition from information sources —various processing heuristics—that overwhelm the syntactic structure that, if built and maintained sufficiently, will lead unfailingly to the correct interpretation. The aim of the present study is to investigate and document one factor that may strengthen the syntactic representation: the prosodic representation. We hypothesized that if readers could be cued to construct a more perceptually salient prosodic representation during silent reading, the corresponding syntactic representation would be "buttressed" against intrusions from heuristics.

Considerable research has demonstrated that syntactic structure is intimately intertwined with prosodic structure (e.g., Bader, 1998; Breen, 2014; Carlson, Clifton, & Frazier, 2001; Cutler, Dahan, & van Donselaar, 1997; Fodor, 2002; Slowiaczak & Clifton, 1980; Speer, Kjelgaard, & Dobroth, 1996; Watson & Gibson, 2005). In particular, Fodor's (2002) implicit prosody hypothesis proposes that during normal silent reading, a default prosodic contour is projected onto the stimulus, which reflects, all things being equal, the most natural prosody for the structure. With respect to reanalysis of difficult syntactic structures, Bader (1998) proposed the prosodic constraint on reanalysis, which states that syntactic reanalysis is more difficult if the prosody must be reanalysed simultaneously. As discussed at length in Breen (2014), these results suggest that suprasegmental or prosodic phrasing provides information about the input's syntactic phrasing, reinforcing the grouping of words into phrasal units (cf. Slowiaczak & Clifton, 1980). The proposal here is that a more conscious instantiation of this suprasegmental phrasing, in the form of auditory perceptual simulation (APS), results in a more salient, resilient prosodic structure and, by association, syntactic structure.

It has also been long recognized that during silent reading, skilled readers automatically generate an "inner voice" (e.g., Breen, 2014; Edfeldt, 1960; Huey, 1968; Rayner, Pollatsek, Ashby, & Clifton, 2012). Consistent with this observation, eye movement data (Ashby & Clifton, 2005) and event-related potential (ERP) data (Ashby, Sanders, & Kingston, 2009) provide evidence that rough phonological representations are included in inner speech during silent reading, as well as evidence that readers create an implicit metrical structure during silent reading (Breen & Clifton, 2011). Under certain conditions, this basic or default prosodic representation can be elaborated, such that it contains richer, more veridical information—that is, a virtual auditory representation of the text approximating the representation that would result from hearing, rather than reading the text. Evidence for this phenomenon, called *auditory perceptual simulation* (APS), is both behavioural and neurological. As far back as 1977, Kosslyn and Matt reported that the prose of a purported "slow speaker" was likely to be read more slowly than the prose of a purported "fast speaker", a result that has since been replicated in several ways (e.g., Alexander & Nygaard, 2008; Drumm & Klin, 2011). Also, specific speakers' voices are simulated under certain conditions during silent reading (Filik & Barber, 2011; Kurby, Magliano, & Rapp, 2009).

These findings have been interpreted within embodied cognition theory, which proposes that semantic meaning is derived from the activation of perceptual features that are associated with physical experiences and sensations (Barsalou, 1999; Barsalou & Hale, 1993; Glenberg & Robertson, 2000; Zwaan, 2004, 2008; Zwaan & Radvansky, 1998). Studies have demonstrated that language processing elicits perceptual simulation: Readers perceptually simulate the motion, action, odour, and other information described in text during reading (Bergen & Wheeler, 2010; Glenberg & Kaschak, 2002; Hubbard, 2010; Louwerse & Connell, 2011; Louwerse & Jeuniaux, 2010; Pecher, Zeelenberg, & Barsalou, 2003, 2004; Pecher & Zwaan, 2005; Solomon & Barsalou, 2004; Zwaan, 2009). Thus, during language processing, readers "see" the depicted objects and events, "smell" the described odour, "act" the described motion, and, importantly for the present study, "hear" the depicted voice in their minds.

More recently, eye-tracking experiments have provided more fine-grained results confirming that not only global reading rates but also discrete eye fixation measures reflect properties of depicted speech in texts. Both Yao and Scheepers (2011) and Stites, Luke, and Christianson (2013) showed that direct quotes, but not indirect quotes, were read more quickly if they were described as having been said by a fast talker (Yao & Scheepers, 2011) or simply as having been said quickly, as in Example 4, from Stites et al.

4. David walked into the room and said quickly (slowly), "I finally found my car keys!"

Yao and Scheepers (2011) used functional magnetic resonance imaging (fMRI) to show a significant boost in brain activity in auditory cortices as people read direct quotes. Assuming that the vast literature on "inner voice" cited above is not wrong, and that basic, default phonological and prosodic representations are indeed generated naturally during silent reading, the additional effects observed in these eye tracking and fMRI studies can be taken as evidence of an elaborated, prosodically richer representation—APS—which can be generated under certain circumstances.

The study reported here is the first to explore APS as a potential strategy to bolster reading comprehension. In particular, we examined whether cueing readers to undertake APS during reading of difficult sentences such as in Example 2 above would reduce the likelihood that readers would end up with "good-enough" interpretations of such sentences. As discussed above, earlier evidence supports the connection between "inner voice" and reading comprehension, showing that suppressing the inner voice hampers comprehension (Slowiaczek & Clifton, 1980) and interferes with readers' ability to notice anomalous words or syntax (Baddeley, Eldridge, & Lewis, 1981). On the other hand, some other work suggests that suppression or lack of inner voice affects only verbatim memory for text (Levy, 1975). No one to our knowledge has examined APS as a more elaborated "inner voice" with respect to its effect on reading comprehension.

In order to investigate the effectiveness of inducing APS to help readers avoid misinterpretation, we developed a novel paradigm. In it, participants were familiarized with the voices of two speakers, one a native American English speaker, and one a non-native American English speaker. These speakers differed in accent and, crucially, speech rate, with the non-native speaker speaking more slowly. Then, prior to silently reading each stimulus sentence, participants were cued with pictures or names of these speakers and were instructed to imagine that the given speaker was saying the given sentence. Participants' eye movements were tracked as they read sentences that were manipulated with respect to their structure (subject-relative clause, SRC, vs. object-relative clause, ORC) and plausibility of thematic roles (plausible, implausible). After each sentence had been read, participants' comprehension was probed using a paraphrase verification task (Kim & Christianson, 2013). An example item in all four conditions, and the paraphrase verification probe (the correct answer to which was always TRUE for the test items), are provided in (5).

5. a. The bird that ate the worm was small. (subject-relative, plausible)
   b. The worm that ate the bird was small. (subject-relative, implausible)
   c. The worm that the bird ate was small. (object-relative, plausible)
   d. The bird that the worm ate was small. (object-relative, implausible)
   e. The bird/worm ate the worm/bird. The bird/worm was small. (T/F)

Three experiments were conducted. In Experiment 1, participants read sentences such as in (5) while their eyes were being tracked without any cue to perform APS. This experiment established baseline normal silent reading rates, patterns, and comprehension for items of this sort. In Experiment 2, participants were cued prior to each sentence to perform APS with pictures of "speakers", and in Experiment 3, participants were cued with recordings of the "speakers" saying their names. A between-subjects design was used to compare APS manipulations across experiments to non-APS normal silent reading. The rationale for this design was based on logistical concerns. If APS versus non-APS reading conditions were included in the same experiment, participants might be likely to maintain whichever reading mode was used first throughout the experiment, and differences between APS reading and non-APS could be washed out. Counterbalancing the order of the reading modes across participants would not have eliminated this concern.

We predicted that the speech rates of the "speakers" would affect the reading speeds of participants who performed APS: Simulating the native speaker would result in faster reading times than simulating the non-native speaker. We also predicted that inducing readers to perform APS would improve comprehension compared to the non-APS baseline, especially in the condition in which the most "good-enough"-driven misinterpretation was predicted to occur—that is, in implausible object-relative sentences. We were not certain how APS would affect reading speed compared to normal silent reading without APS, however.

EXPERIMENT 1

In Experiment 1, we examined how readers read and process sentences such as (5) silently, without being cued to initiate APS. The purpose of this experiment was to establish baselines for reading speed, eye movement patterns, and comprehension.

Method

Participants

Fifty-three native English speakers with normal or corrected-to-normal vision were recruited through the Educational Psychology subject pool at University of Illinois at Urbana-Champaign. These participants did not participate in either of the other two experiments. A survey was used to collect participants' language and education backgrounds, ensuring the relative homogeneity of the sample in these aspects. Participants received either one research credit or $7 for their participation. Thirteen participants' data were excluded due to experimental errors. The data from the remaining 40 participants were analysed.

Materials

Forty-eight target sentences were used in Experiment 1 (and the same sentences were used in Experiments 2–3, see the Appendix for the target sentences). One hundred and thirty filler sentences were included, consisting of a mix of complex and simple sentences, all of which were grammatical. As described above, the target sentences were manipulated with respect to structure and plausibility, resulting in a 2 (subject-relative, object-relative) × 2 (plausible, implausible) fully crossed, within-participants and within-items design. Materials were distributed pseudorandomly for each participant across four lists according to a

Latin square, so that every participant saw each item only once.

One half of the items were converted from the passives used in Ferreira (2003) and Christianson et al. (2010) into relative clauses by Lim and Christianson (2013b). The other half were created by Lim and Christianson. Plausibility of thematic role assignment, for example whether it is more likely for a bird to eat a worm or a worm to eat a bird, was normed by Ferreira (2003) for half the items and by Lim and Christianson (2013b) for the other half of the items.

There is a large psycholinguistic literature detailing the increased processing difficulty of object-relative clauses compared to subject-relative clauses and several competing theories attempting to account for this processing asymmetry, none of which are relevant to the current investigation (e.g., De Villiers, Tager, Flusberg, Hakuta, & Cohenl, 1979; Gennari & MacDonald, 2008; Gibson, 1998; Gordon, Hendrick, & Johnson, 2004; Gordon, Hendrick, Johnson, & Lee, 2006; Gordon, Hendrick, & Levine, 2002; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Traxler, Morris, & Seely, 2002; Van Dyke & McElree, 2006; Weckerly & Kutas, 1999; Wells, Christiansen, Race, Acheson, & MacDonald, 2009). What is relevant, however, is the novel finding by Ferreira (2003) that syntactic structures that are more difficult to process—described as having noncanonical word order—are likely to be misinterpreted, especially when those structures communicate meanings that are implausible. For example, Ferreira showed that native English speakers are more likely to misinterpret implausible passive sentences like *The bird was eaten by the worm* than plausible passive sentences like *The worm was eaten by the bird* or implausible active sentences like *The worm ate the bird*. We therefore predicted that reading speeds, eye movement patterns, and comprehension accuracy should all be most impacted in the implausible object-relative condition. Structure and plausibility should have independently observable effects in plausible object-relatives and implausible subject-relatives. The plausible subject-relative condition should be the easiest by all measures and served as the baseline for comparison.

Apparatus

Eye movement data and offline comprehension data were collected using an SR Research EyeLink 1000 remote desktop eye tracker running SR Research Experiment Builder software. A chin rest and forehead rest were used to stabilize participants' heads. Sentences were presented in black Courier New monotype font (14 pt) on a white background on the computer monitor, which was approximately 70 cm away from the participants. At this distance, 1° of visual angle subtended approximately 3 characters. All viewing was binocular, but data were recorded from only the right eye.

Procedure

The entire experimental session lasted on average 50 min. Participants provided informed consent and were then calibrated on the eye tracker using a nine-point calibration procedure. After a six-item practice session, participants were asked to read each sentence silently before completing a paraphrase verification task. After the eye-tracking portion of the experiment, the social attractiveness survey was administered on a different computer in the same room.

Results

Fixations and sentence reading time data were trimmed before data analysis. Fixation times shorter than 80 ms and longer than 1200 ms were excluded. Sentence reading times were log transformed and scaled to reduce collinearity. Remaining data points were examined and trimmed within each condition: Reading times greater or smaller than two standard deviations from the mean of that particular condition were deleted (approximately 3.6% of the data in total). Reading time data for the entire sentence and four specific regions of interest were analysed using linear mixed-effects models (Baayen, 2008); logit mixed models were applied to examine binomial response accuracy data (Jaeger, 2008). All

Table 1. *Means and standard deviations of sentence reading times and response accuracy in three experiments*

			Sentence reading time		*Accuracy*	
Experiment	*"Speaker"*	*Structure*	*Plausible*	*Implausible*	*Plausible*	*Implausible*
1	No APS	Subject-RC	4071.77 (2124.40)	4679.54 (2525.83)	.94 (.24)	.85 (.36)
		Object-RC	4376.42 (2290.88)	4950.72 (2679.64)	.86 (.35)	.58 (.49)
2	Native	Subject-RC	3576.54 (1924.85)	3864.53 (2194.14)	.96 (.19)	.95 (.22)
		Object-RC	4196.10 (2244.21)	4603.20 (2758.85)	.88 (.32)	.70 (.46)
	Non-native	Subject-RC	3933.12 (2647.61)	4010.09 (2060.25)	.95 (.21)	.92 (.28)
		Object-RC	4546.79 (2528.96)	4800.14 (2850.15)	.87 (.34)	.69 (.46)
3	Native	Subject-RC	3884.62 (1934.45)	4100.96 (2322.18)	.94 (.23)	.96 (.20)
		Object-RC	4226.56 (2168.60)	4529.98 (2478.67)	.90 (.30)	.73 (.45)
	Non-native	Subject-RC	3970.37 (1968.32)	4124.54 (1882.06)	.95 (.22)	.94 (.23)
		Object-RC	4482.30 (2311.70)	4642.13 (2348.27)	.88 (.33)	.74 (.44)

Note: Reading time in ms. APS = auditory perceptual simulation; RC = relative clause.

analyses were performed using maximal random effects structures (Barr, Levy, Scheepers, & Tily, 2013)—that is, with random slopes and intercepts for participants and items. These maximal random effects models were compared to models with only random intercepts; in the interest of space, we report below (for all experiments) the full maximal random effects models only in cases in which the addition of the random slopes changed the results.

Sentence reading time

A stepwise selection procedure was applied to determine the best fitted linear mixed-effects model, and only factors that were significant or nearly so, or which participated in interactions, were included in the final, best fitted model. Model output tables for all analyses in all experiments can be found in the online Supplemental Material. The best fitted model revealed that syntactic structure ($t = -3.26$, $p < .001$), plausibility ($t = -4.54$, $p < .001$), and trial presentation order ($t = -6.14$, $p < .001$) had significant effects on sentence reading time without any interactions; subjects and items were significant random intercepts and slopes in the model. Trial presentation order was included to account for changes in reading speed as participants progressed through the experiment (i.e., practice effects). Participants read plausible sentences faster than implausible sentences and read SRCs faster than ORCs, as expected. Across the experiment, participants read sentences that were presented later faster than those presented earlier. Descriptive statistics for sentence reading time are presented in Table 1.

Interest area analyses

The four specific interest areas were: the first noun (FN) region, which was the subject noun of the main clause; the relative clause verb region (RCV); the second noun (SN) region, which was the noun within the relative clause; and the main clause (MC) region, consisting of the remainder of the sentence. Reading times for the words within the MC region were summed to create a single dependent variable. An example of the specific regions is presented in Example 6.

6. The|bird| that|ate| the|worm| was small.
FN RCV SN MC

Seven eye movement measures were included in the interest area analyses. *First-fixation duration* is the duration of the first fixation made in the target region. *Gaze duration* is the summed duration of all fixations in the target region from when it is first fixated until the eyes move to another region in either direction. When there is only one fixation on a region, first-fixation duration equals gaze duration. *Go-past time*, also called regression path duration, is the total time of all fixations in a region from when it is first entered until it is exited to the right (or until the button is pressed to move on to the comprehension probe, in the case of the MC region), including fixations on previous regions that are refixated after regressive saccades. *Total reading time* is the summed duration of all fixations in the target region during the trial. *Regressions-in* reflects whether a regression was made into the target region from a later part of the sentence. *Regressions-out* reflects whether a regression was made out of the target region to an earlier part of the sentence. *Skipping* shows whether the region was fixated or not. The four duration variables are recorded and reported descriptively in milliseconds, while the other three are binomial (1 if a regression in/out was made or a region was skipped, 0 if otherwise). In the first noun region, we only examined gaze duration, total reading time, and regressions-in. In the other three regions, we examined all seven measures. See the descriptive data in Table 2.

In the first noun region, plausibility was a significant predictor of total reading time ($t = -4.81$, $p < .001$). Implausible sentences led to longer total reading times than plausible ones on the first noun. No factors were significant in either gaze duration or regressions-in.

In the relative clause verb region, structure significantly influenced first-fixation durations and gaze durations and marginally affected go-past times. Readers had longer first-fixation durations ($t = -7.17$, $p < .001$), gaze durations ($t = -4.54$, $p < .001$), and go-past times ($t = -1.82$, $p < .1$) in ORCs than in SRCs. Total reading times were significantly influenced by plausibility ($t = -2.52$, $p < .01$) and structure ($t = -4.01$, $p < .001$) without any interaction (ORC > SRC, implausible > plausible). Both conditions marginally affected regressions-in (plausibility: $z = -1.8$, $p < .1$; structure: $z = -1.77$, $p < .1$) and regressions-out (plausibility: $z = -1.8$, $p < .1$; structure: $z = -1.77$, $p < .1$). Only structure ($z = 3.25$, $p < .01$) significantly affected skipping on this region, with more skipping in SRCs.

In the second noun region, plausibility significantly predicted first-fixation duration ($t = -2.21$, $p < .05$), such that readers spent less time reading this region when the sentence was plausible. Plausibility also significantly influenced gaze durations ($t = -2.25$, $p < .05$) and go-past times ($t = -2.35$, $p < .05$). Plausible sentences yielded shorter gaze durations and faster go-past times. Both structure ($t = -5.26$, $p < .001$) and plausibility ($t = -2.50$, $p < .05$) significantly affected total reading times in the same directions. Logit mixed-effects models demonstrated that plausibility was a significant predictor of regressions-in ($z = -2.18$, $p < .05$), regressions-out ($z = -2.18$, $p < .05$), and skipping rate ($z = 2.67$, $p < .05$) on the second noun region. Readers were less likely to make regressions-out or -in, but more likely to skip the second noun, in plausible sentences.

In the main clause region, structure ($t = -2.15$, $p < .05$) and plausibility ($t = -2.31$, $p < .05$) influenced go-past times with no interaction. Plausibility ($t = -2.04$, $p < .05$) was a significant predictor of gaze durations. Plausibility also affected total reading times ($t = -4.14$, $p < .001$), which were inflated in implausible sentences. There were no other significant effects on this region.

Response accuracy data

The means and standard deviations of response accuracy in Experiment 1 are presented in Table 1. Using the same model selection procedure and criteria as those with the linear mixed models, the final best fitted logit model included structure ($z = 7.12$, $p < .001$), plausibility ($z = 7.17$, $p < .001$), and trial presentation order ($z = 3.37$, $p < .001$) as main effects and random intercepts and slopes for items and participants. There was a two-way interaction between structure and plausibility ($z = -3.18$, $p < .001$); readers were most likely to misinterpret sentences when they were implausible ORCs. Participants were more likely to interpret items accurately later in the experiment.

Table 2. *Means and standard deviations of first-fixation duration, gaze duration, go-past time, total times, regressions-in, regressions-out, and skipping on four interest areas in three experiments*

Expt	*Region*	*Speaker*	*Structure*	*Plausibility*	*FFD*	*GD*	*GP*	*TT*	*RI*	*RO*	*SK*
1	FN	N/A	Subject-RC	plaus	—	302.87 (224.91)	—	501.32 (436.61)	.58 (.49)	—	—
				implaus	—	303.37 (197.17)	—	611.60 (523.88)	.70 (.46)	—	—
			Object-RC	plaus	—	287.79 (171.60)	—	548.35 (477.20)	.65 (.48)	—	—
				implaus	—	299.91 (192.50)	—	612.07 (493.08)	.70 (.46)	—	—
	RCV		Subject-RC	plaus	237.01 (96.81)	284.17 (151.40)	377.83 (247.96)	517.73 (423.85)	.33 (.47)	.21 (.41)	.14 (.34)
				implaus	225.87 (95.37)	279.08 (167.00)	408.41 (376.10)	606.42 (480.76)	.43 (.50)	.23 (.42)	.15 (.35)
			Object-RC	plaus	251.01 (103.99)	325.36 (197.35)	454.13 (453.90)	599.29 (468.67)	.32 (.47)	.15 (.35)	.10 (.30)
				implaus	252.29 (98.24)	319.83 (172.87)	479.24 (616.44)	675.17 (493.39)	.29 (.45)	.16 (.37)	.09 (.29)
	SN		Subject-RC	plaus	224.95 (79.43)	265.89 (136.31)	370.81 (348.67)	444.56 (400.91)	.29 (.45)	.14 (.35)	.20 (.40)
				implaus	232.98 (85.64)	290.95 (156.66)	430.89 (495.28)	543.93 (455.22)	.29 (.45)	.19 (.40)	.15 (.36)
			Object-RC	plaus	220.19 (79.75)	265.10 (141.93)	397.24 (360.70)	494.22 (414.18)	.33 (.47)	.20 (.40)	.18 (.38)
				implaus	220.21 (81.99)	257.6 9 (135.06)	372.27 (304.78)	584.15 (523.41)	.38 (.49)	.19 (.39)	.14 (.35)
	MC		Subject-RC	plaus	268.03 (134.18)	796.73 (480.06)	2030.02 (1468.39)	1158.30 (653.80)	.00	.69 (.46)	.02 (.13)
				implaus	277.36 (147.01)	727.23 (412.76)	2454.48 (1793.78)	1267.82 (766.49)	.00	.80 (.40)	.02 (.14)
			Object-RC	plaus	269.02 (136.88)	747.4 (493.36)	2213.54 (1648.29)	1148.64 (738.72)	.00	.76 (.43)	.03 (.16)
				implaus	280.23 (163.03)	719.00 (440.15)	2600.80 (1980.00)	1268.25 (825.71)	.00	.77 (.42)	.01 (.11)
2	FN	Native	Subject-RC	plaus	—	275.37 (142.06)	—	467.32 (390.32)	.57 (.50)	—	—
				implaus	—	275.00 (158.53)	—	499.34 (417.10)	.65 (.48)	—	—
			Object-RC	plaus	—	267.86 (158.08)	—	542.21 (440.69)	.70 (.46)	—	—
				implaus	—	264.40 (148.81)	—	560.92 (509.98)	.69 (.46)	—	—
		Non-native	Subject-RC	plaus	—	284.69 (166.19)	—	508.31 (438.29)	.61 (.49)	—	—
				implaus	—	279.95 (162.47)	—	532.22 (420.52)	.66 (.48)	—	—
			Object-RC	plaus	—	294.42 (206.29)	—	644.34 (569.16)	.73 (.44)	—	—
				implaus	—	277.16 (153.15)	—	609.00 (488.71)	.70 (.46)	—	—
	RCV	Native	Subject-RC	plaus	216.29 (79.99)	251.85 (128.33)	357.99 (269.61)	435.85 (378.59)	.35 (.48)	.22 (.41)	.22 (.41)
				implaus	217.66 (81.94)	259.37 (127.12)	341.78 (236.38)	494.64 (437.51)	.39 (.49)	.18 (.38)	.19 (.39)
			Object-RC	plaus	234.22 (89.28)	286.14 (140.44)	407.59 (378.75)	514.84 (358.46)	.27 (.45)	.16 (.37)	.11 (.32)
				implaus	238.76 (101.27)	283.06 (143.06)	411.72 (512.74)	575.32 (459.01)	.31 (.46)	.14 (.35)	.10 (.29)
		Non-native	Subject-RC	plaus	219.38 (89.60)	268.8 (142.22)	386.55 (292.45)	517.37 (421.58)	.37 (.48)	.20 (.40)	.16 (.37)
				implaus	219.62 (88.89)	262.52 (135.13)	356.40 (304.97)	504.72 (392.30)	.40 (.49)	.17 (.37)	.16 (.37)
			Object-RC	plaus	246.19 (100.40)	300.15 (160.03)	413.97 (420.41)	580.87 (418.01)	.32 (.47)	.14 (.35)	.10 (.31)
				implaus	242.02 (89.75)	302.99 (179.70)	401.42 (342.87)	627.35 (486.14)	.36 (.48)	.14 (.35)	.14 (.34)

	SN	Native	Subject-RC	plaus	219 (83.86)	254.77 (123.32)	254.77 (123.32)	406.35 (334.69)	.25 (.43)	.25 (.43)	.14 (.35)
				implaus	223.83 (88.48)	260.35 (122.74)	260.35 (122.74)	450.39 (374.41)	.28 (.45)	.28 (.45)	.17 (.37)
			Object-RC	plaus	210.87 (80.00)	239.4 (116.18)	239.4 (116.18)	450.81 (396.23)	.32 (.47)	.32 (.47)	.20 (.40)
				implaus	211.74 (83.16)	238.86 (118.63)	238.86 (118.63)	532.35 (532.07)	.33 (.47)	.33 (.47)	.19 (.39)
		Non-native	Subject-RC	plaus	226.56 (88.20)	260.48 (128.42)	260.48 (128.42)	420.30 (351.60)	.25 (.43)	.25 (.43)	.18 (.39)
				implaus	226.36 (81.78)	261.04 (128.73)	261.04 (128.73)	437.85 (343.63)	.23 (.42)	.23 (.42)	.16 (.37)
			Object-RC	plaus	213.91 (84.93)	239.31 (116.29)	239.31 (116.29)	484.09 ((449.11)	.35 (.48)	.35 (.48)	.21 (.41)
				implaus	215.02 (94.53)	244.76 (141.67)	244.76 (141.67)	519.26 (440.69)	.37 (.48)	.37 (.48)	.20 (.40)
	MC	Native	Subject-RC	plaus	238.19 (104.79)	779.34 (479.12)	1878.15 (1374.34)	1088.18 (647.85)	.00	.73 (.44)	.00 (.06)
				implaus	233.91 (86.60)	741.79 (484.83)	2155.13 (1632.20)	1154.59 (704.83)	.00	.81 (.39)	.02 (.14)
			Object-RC	plaus	244.69 (116.11)	732.19 (507.92)	2230.53 (1632.53)	1159.61 (746.96)	.00	.82 (.38)	.01 (.10)
				implaus	243.75 (118.44)	706.33 (455.97)	2489.99 (2011.95)	1179.43 (720.92)	.00	.83 (.37)	.01 (.08)
		Non-native	Subject-RC	plaus	236.37 (94.93)	790.30 (518.91)	2096.33 (1549.82)	1180.38 (801.48)	.00	.79 (.41)	.01 (.12)
				implaus	238.03 (98.73)	785.94 (546.26)	2129.51 (1434.96)	1168.66 (762.38)	.00	.80 (.40)	.01 (.12)
			Object-RC	plaus	248.38 (110.87)	767.77 (518.53)	2490.54 (1769.75)	1223.15 (799.41)	.00	.85 (.36)	.01 (.12)
				implaus	246.53 (116.71)	718.64 (470.74)	2703.80 (2025.52)	1268.80 (856.68)	.00	.85 (.35)	.01 (.10)
3	FN	Native	Subject-RC	plaus	—	334.17 (293.23)	—	477.25 (504.10)	.30 (.46)	—	—
				implaus	—	329.48 (294.37)	—	489.83 (493.40)	.31 (.46)	—	—
			Object-RC	plaus	—	357.52 (307.73)	—	594.07 (536.38)	.33 (.47)	—	—
				implaus	—	357.54 (305.67)	—	634.11 (583.07)	.34 (.47)	—	—
		Non-native	Subject-RC	plaus	—	342.07 (319.82)	—	482.89 (512.52)	.29 (.46)	—	—
				implaus	—	335.93 (302.98)	—	501.71 (494.34)	.32 (.47)	—	—
			Object-RC	plaus	—	365.97 (322.77)	—	630.26 (572.35)	.33 (.47)	—	—
				implaus	—	366.22 (324.07)	—	653.24 (584.74)	.35 (.48)	—	—
	RCV	Native	Subject-RC	plaus	216.43 (83.44)	334.17 (293.23)	327.18 (212.13)	477.25 (504.10)	.36 (.48)	.17 (.37)	.20 (.40)
				implaus	216.30 (84.00)	329.48 (294.37)	332.54 (234.26)	489.83 (493.40)	.45 (.50)	.17 (.38)	.20 (.40)
			Object-RC	plaus	219.09 (89.68)	357.52 (307.73)	374.18 (268.17)	594.07 (536.38)	.28 (.45)	.13 (.34)	.11 (.32)
				implaus	222.16 (91.11)	357.54 (305.67)	373.11 (314.00)	634.11 (583.07)	.28 (.45)	.11 (.31)	.11 (.31)
		Non-native	Subject-RC	plaus	216.08 (82.68)	342.07 (319.82)	359.37 (263.87)	482.89 (512.52)	.36 (.48)	.19 (.39)	.19 (.39)
				implaus	219.59 (87.05)	335.93 (302.98)	370.09 (288.42)	501.71 (494.34)	.45 (.50)	.20 (.40)	.19 (.39)
			Object-RC	plaus	222.45 (88.41)	365.97 (322.77)	381.20 (308.34)	630.26 (572.35)	.31 (.46)	.13 (.34)	.12 (.32)
				implaus	224.64 (96.23)	366.22 (324.07)	431.33 (416.00)	653.24 (584.74)	.33 (.47)	.16 (.37)	.10 (.30)

(*Continued overleaf*)

Table 2. *Continued*

Expt	*Region*	*Speaker*	*Structure*	*Plausibility*	*FFD*	*GD*	*GP*	*TT*	*RI*	*RO*	*SK*
	SN	Native	Subject-RC	plaus	215.90 (84.72)	246.68 (133.30)	306.39 (238.65)	371.46 (317.48)	.26 (.44)	.13 (.33)	.20 (.40)
				implaus	224.18 (87.92)	260.53 (126.47)	340.42 (273.89)	414.71 (330.45)	.28 (.45)	.13 (.34)	.19 (.39)
			Object-RC	plaus	207.68 (68.73)	235.60 (116.43)	343.00 (278.33)	428.02 (377.67)	.36 (.48)	.20 (.40)	.21 (.41)
				implaus	202.76 (70.47)	228.89 (102.80)	332.20 (266.70)	453.22 (417.31)	.36 (.48)	.17 (.38)	.24 (.43)
		Non-native	Subject-RC	plaus	211.35 (79.10)	245.40 (122.89)	320.69 (286.86)	390.46 (324.50)	.26 (.44)	.12 (.33)	.19 (.39)
				implaus	226.65 (93.73)	265.57 (143.65)	359.03 (310.79)	416.77 (306.91)	.29 (.45)	.17 (.37)	.19 (.39)
			Object-RC	plaus	217.34 (82.91)	247.81 (121.74)	357.79 (287.73)	457.38 (420.83)	.37 (.48)	.19 (.39)	.26 (.44)
				implaus	211.52 (104.34)	246.63 (154.31)	355.24 (315.84)	486.46 (466.98)	.38 (.49)	.20 (.40)	.20 (.40)
	MC	Native	Subject-RC	plaus	240.24 (101.58)	784.76 (446.47)	2020.84 (1439.80)	1184.08 (763.30)	.00	.79 (.41)	.01 (.08)
				implaus	247.03 (107.08)	765.25 (489.10)	2098.86 (1459.36)	1193.75 (708.62)	.00	.81 (.40)	.00 (.06)
			Object-RC	plaus	247.94 (122.15)	763.77 (475.85)	2224.33 (1571.36)	1190.45 (698.24)	.00	.83 (.38)	.01 (.09)
				implaus	253.26 (122.15)	744.92 (483.50)	2406.14 (1890.92)	1180.19 (714.78)	.00	.83 (.38)	.00 (.06)
		Non-native	Subject-RC	plaus	242.92 (102.36)	812.41 (527.39)	2033.68 (1467.43)	1193.44 (799.43)	.00	.80 (.40)	.01 (.10)
				implaus	245.73 (111.85)	788.91 (505.09)	2141.97 (1398.12)	1212.98 (705.02)	.00	.83 (.38)	.01 (.10)
			Object-RC	plaus	245.05 (103.15)	770.12 (527.61)	2365.65 (1724.80)	1234.42 (813.72)	.00	.83 (.38)	.01 (.12)
				implaus	257.92 (126.44)	764.53 (531.91)	2425.81 (1784.77)	1249.99 (743.32)	.00	.85 (.36)	.00 (.05)

Note: Standard deviations in parentheses. Expt = experiment; FFD = first-fixation duration; GD = gaze duration; GP = go-past time; TT = total times; RI = regressions-in; RO = regressions-out; SK = skipping; FN = first noun; RCV = relative clause verb; SN = second noun; MC = main clause; RC = relative clause; plaus = plausible; implaus = implausible.

EXPERIMENT 2

Experiment 2 was designed to test the hypothesis that APS would improve comprehension accuracy of sentences in which semantic plausibility information was most likely to interfere with syntactic structure building (i.e., implausible object-relatives), which were observed in Experiment 1 to pose the most problems for interpretation, consistent with previous work. We reasoned that APS should buttress the syntactic representation with an elaborated prosodic representation, reducing interferential effects of the semantics. A novel paradigm was developed, in which participants in a silent reading eye-tracking experiment were shown a picture of one of two "speakers". It was hypothesized that these pictures, in combination with initial instructions asking participants to imagine the voices of the "speakers" as they read, would serve as effective prompts for APS.

Method

Participants

Ninety-eight native English speakers with normal or corrected-to-normal vision were recruited through the Educational Psychology subject pool at University of Illinois at Urbana-Champaign. These participants did not participate in either of the other two experiments. Participants received either one research credit or $7 for their participation. Eight participants' data were not recorded completely due to experimenter error, and 10 participants' data were excluded because their experimental sessions were disrupted. The data from the remaining 80 participants were analysed.

Materials

In order to induce participants to perceptually simulate the speech of distinct speakers, we provided them with examples of both a native and a non-native English speaker's voice. Four 500-word passages were used as the texts for the speech. These narrative passages, balanced for difficulty (Flesch–Kinkaid reading level: Text 1 = 11.8, Text 2 = 12.2, Text 3 = 12.5, Text 4 = 12), were selected from *National Geographic* on topics such as organic fruit. One female native American English speaker and one female Chinese non-native American English speaker of approximately the same age (early 20s) were recruited to read these passages. Reading speeds were compared using a paired *t*-test, confirming that the native English speaker read the passages significantly faster than the non-native English speaker, $t(3) = -72.38$, $p < .001$.[1]

The photos of the speakers were not of the actual speakers, but rather stock, noncopyrighted photos found on the Internet and controlled as closely as possible for age, socioeconomic status, and expression. The photo of the native speaker was of a blonde, Caucasian woman who appeared to be in her early 30s. The photo of the non-native English speaker was of a Chinese woman who appeared to be in her early 30s. The pictures showed the women from the shoulders up, and both women were dressed in a business jacket and blouse.

The same target sentences as those in Experiment 1 and 96 filler sentences were used in Experiment 2. All sentences had quotation marks to remind participants that they should read them as direct quotations (Stites et al., 2013; Yao & Scheepers, 2011). In addition to the reading task, a social attractiveness scale questionnaire was used to measure participants' attitudes towards the native and non-native speakers and speech. Based on previous studies (Callen, Callois, & Forbes, 1983; Edwards, 1977; Gass & Varonis, 1984; Giles, 1972; Vornik, Sharman, & Garry, 2003), 14 factors were included in the questionnaire, such as comprehensibility, accent, and trustworthiness. (See online Supplemental Material for the complete survey.) Participants rated each speaker on the 14 attributes on a 1–7 Likert scale. The social attractiveness data were collected to ensure that sociolinguistic biases against speakers/accents were not the source of the effects that were observed. Each participant filled out the questionnaire after the eye-tracking portion of the experiment.

[1]The time duration for native speech: Text 1 = 2 min 50 s; Text 2 = 2 min 57 s; Text 3 = 2 min 49 s; Text 4 = 2 min 57 s. The time duration for non-native speech: Text 1 = 4 min 08 s; Text 2 = 4 min 18 s; Text 3 = 4 min 08 s; Text 4 = 4 min 20 s.

Apparatus

The same equipment as that in Experiment 1 was used in Experiment 2. A separate desktop computer in the same room was used to collect the survey results.

Procedure

The entire experimental session lasted on average 50 minutes. Calibration followed the same procedure as that in Experiment 1. Next, the participant listened to two recordings, one by the native English speaker, the other by the non-native English speaker, while the corresponding photograph was displayed on the computer screen. The order of recordings was counterbalanced across participants, with each one presented first half of the time. Participants then read and responded to the six practice items and, after any questions, proceeded to read and respond to the experiment items. One or the other "speaker" photograph was randomly presented an equal number of times prior to each sentence read in the experiment to cue the "speaker's" voice as participants read the sentences. One concern was that participants' memories for the speakers' voices would fade as the experiment progressed. To address this concern, another recording set of the native and non-native speaker reading similar texts aloud was played halfway through the experiment to remind participants what they sounded like. After the second reading portion of the experiment, the two social attractiveness scale questionnaires (one for each speaker) were administered to measure participants' attitudes towards the speakers.

Results

Following the same data trimming procedure as that in Experiment 1, approximately 3.5% of the data were removed prior to analysis.

Sentence reading time

The same procedure was applied to determine the best fitted linear mixed-effects model. The final model revealed significant effects of structure ($t = -8.59$, $p < .001$), plausibility ($t = -4.08$, $p < .001$), "speaker" ($t = -5.55$, $p < .001$), and trial presentation order ($t = -5.53$, $p < .001$) on total sentence reading time without any interactions. Participants took longer to read ORCs than SRCs. Implausible sentences were read more slowly than plausible sentences. Sentences attributed to the non-native "speaker" yielded longer reading times than sentences attributed to the native "speaker". Regardless of structure, plausibility, or "speaker", participants read sentences presented later faster than sentences presented earlier.

To ensure that participants' memories for the "speakers'" voices did not fade over time, we presented the reading materials in two blocks, separated by a re-presentation of another voice sample for each speaker. Interestingly, and consistent with APS effects, we found that the main effect of "speaker" was lacking in post hoc subanalyses of the 12 items immediately preceding the re-presentation, but reemerged in the 12 items immediately following the re-presentation. This pattern was also observed in Experiment 3. The result strongly supports the claim that readers were simulating the voices as they read. See "auditory re-representation effect in Experiment 2" in Supplemental Material for more information.

Interest area analysis

In the first noun region, "speaker" condition had a significant effect on gaze duration ($t = -2.8$, $p < .01$). Readers processed the first noun much faster when the target sentence was presented as having been said by the native speaker. Structure ($t = -3.67$, $p < .001$) and "speaker" conditions ($t = -4.62$, $p < .001$) significantly affected total reading time without any interaction. Total reading times were longer in the ORC condition than in SRC. Total reading time was, overall, longer when participants were cued by the non-native speaker's photo. Structure also significantly affected regressions into the first noun region ($z = -5.56$, $p < .001$). Participants made more regressions back to the first noun in ORCs (70.47%) than they did in SRCs (62.28%).

In the relative clause verb region, first-fixation durations were significantly inflated in ORCs ($t = -7.40$, $p < .001$) and following the non-native "speaker" photo ($t = -1.83$, $p < .1$). ORC structure also yielded inflated gaze durations ($t = -4.97$, $p < .001$) and go-past times ($t = -3.44$, $p < .001$). Total reading time times on the verb were also significantly greater in ORCs ($t = -4.05$, $p < .001$) and following the non-native "speaker's" photo ($t = -4.58$, $p < .001$). The plausible condition ($t = -1.79$, $p < .1$) led to marginally shorter total reading times in this region. All three factors, syntactic structure ($z = -2.15$, $p < .05$), plausibility ($z = 3.48$, $p < .001$), and "speaker" ($z = -2.13$, $p < .05$), had significant effects on regressions into the relative clause verb. Participants made more regressions into the verb region in SRCs and when sentences were implausible. The non-native "speaker" also triggered more regressions into this region. Plausibility marginally predicted regressions-out ($z = -1.86$, $p < .1$), while syntactic structure significantly affected regressions-out ($z = 3.5$, $p < .001$). Participants made more regressions-out in plausible sentences than in implausible ones, and fewer in ORCs than in SRCs. For skipping, only syntactic structure ($z = 4.58$, $p < .001$) was a significant predictor; more skipping of the relative clause verb occurred in SRCs.

In the second noun region, only sentence structure significantly affected first-fixation durations ($t = 4.05$, $p < .001$), gaze durations ($t = 3.07$, $p < .001$), and total fixation times ($t = -2.81$, $p < .001$). First-fixation durations and gaze durations on the second noun were significantly longer in SRCs. However, total reading times were shorter in SRCs than in ORCs. Plausibility also significantly influenced total reading times ($t = -1.85$, $p < .05$). Plausible sentences yielded shorter total reading times than implausible sentences. There were significant structure effects on regressions-in ($z = -4.96$, $p < .001$) and regressions-out ($z = -7.22$, $p < .001$). Readers made fewer regressions into and out of the second noun in SRCs, demonstrating again the overall greater difficulty of processing ORCs. Structure also significantly affected skipping ($z = -2.16$, $p < .05$). Readers skipped this region more often in ORCs. ORC sentences yielded longer first-fixation durations in the MC region than SRC sentences ($t = -2.56$, $p < .01$). There was a "speaker" effect on total reading time ($t = -3.44$, $p < .001$), such that the native "speaker" condition yielded significantly shorter total times. There were no significant effects on gaze duration, go-past time, regressions-in, regressions-out, or skipping.

Response accuracy data

Using the same model selection procedure and criteria as those with the linear mixed models, the final best fit logit model included structure ($z = 9.53$, $p < .001$), plausibility ($z = 6.66$, $p < .001$), trial order ($z = 4.46$, $p < .001$), and the interaction of structure and plausibility ($z = -3.90$, $p < .001$) as main effects and random intercepts and slopes for participants and items. Structure, plausibility, and trial order had significant effects on response accuracy with no interactions. Participants made more errors in the ORC condition than in the SRC condition. They were more likely to misinterpret implausible sentences than plausible sentences. Comprehension improved as the experiment progressed. "Speaker" did not have significant effect on accuracy. There was only a 1% difference in accuracy between APS of native and APS of non-native speech.

See the online Supplemental Material and Zhou and Christianson (2015) for results and discussion of the social attractiveness survey results.

EXPERIMENT 3

Experiment 2 showed that the native and non-native speakers' photos were effective cues for readers to perceptually simulate native and non-native speech rate during silent reading. If the photos did indeed trigger APS, we reasoned that auditory cues should also be effective at triggering readers to perceptually simulate depicted speakers' voices. By way of replication and exploration of the APS effect, in Experiment 3 we replaced the photos of the speakers with recordings of them saying their names. We predicted that these recordings would also prompt APS and drive similar reading patterns to those observed in Experiment 2.

Method

Participants

Ninety-seven native English speakers with normal or corrected-to-normal vision and who did not participate in Experiment 1 or 2 were recruited through the Educational Psychology subject pool at University of Illinois at Urbana-Champaign. Participants received either one research credit or $7 for their participation. Seventeen participants' data were excluded due to bad calibration or excessive drifts. The data from the remaining 80 participants were analysed.

Materials

The materials were identical to those of Experiment 2, except that the speakers' photos were not used as APS prompts prior to the presentation of each sentence. Instead, two recordings of the speakers saying their names ("Susan" or "Xiaofu") were used.

Apparatus

The same equipment as that in Experiment 1 was used in Experiment 3.

Procedure

All procedures were identical to those of Experiment 2.

Results

Following the same data trimming procedure as that in Experiments 1 and 2, approximately 4.2% of the data were removed prior to analysis.

Sentence reading time

The same model-fitting procedure as that in Experiments 1 and 2 was used, and the best fitted model for total sentence reading time included plausibility ($t = -4.52$, $p < .001$), structure ($t = -2.93$, $p < .001$), "speaker" ($t = -3.25$, $p < .001$), and trial order ($t = -7.23$, $p < .001$), along with random slopes and intercepts for participants and items. Plausible sentences yielded significantly shorter total reading times than implausible sentences. SRCs required less time to read than ORCs. Participants read sentences preceded by the native speaker's name faster than sentences preceded by the non-native speaker's name.[2] Items presented later in the experiment were read faster than those presented earlier.

Interest area analysis

The same interest areas as those in Experiments 1 and 2 were analysed in Experiment 3 for the same measures.

In the first noun region, linear mixed-effects results revealed that "speaker" significantly influenced gaze durations ($t = -2.27$, $p < .05$) on the first noun. When the sentence was preceded by the native speaker's name, gaze durations were shorter on the first noun. Syntactic structure significantly affected the total reading times ($t = -2.17$, $p < .05$) on the first noun, such that the first noun was read more slowly in ORCs than in SRCs. There were no effects on regressions into this region.

In the relative clause verb region, structure significantly affected first-fixation durations ($t = -7.17$, $p < .001$), gaze durations ($t = -4.54$, $p < .001$), and regressions-out ($z = -3.77$, $p < .001$). First-fixation durations and gaze durations on the relative clause verb were inflated in ORCs compared to SRCs, and more regressions-out of the verb were initiated in ORCs. Structure marginally influenced go-past times ($t = -1.82$, $p < .1$). Total reading times were significantly influenced by structure ($t = -4.01$, $p < .001$) and plausibility ($t = -2.52$, $p < .001$). No interactions were significant. Structure significantly influenced the regressions-out of this region. None of the factors significantly impacted regressions into the relative clause verb region or skipping.

In the second noun region, syntactic structure and "speaker" significantly influenced first-fixation

[2]We also examined the auditory re-presentation effect in Experiment 3. The results replicated the patterns in Experiment 2. Participants' memory of the "speakers" gradually faded away as the time went on and then returned after re-presentation. See "auditory re-representation effect in Experiment 3" in the Supplemental Material for details.

durations (structure: $t = 2.01$, $p < .05$; "speaker": $t = 2.00$, $p = .05$) and total times (structure: $t = -2.85$, $p < .001$; "speaker": $t = -2.05$, $p < .05$), and marginally affected the gaze duration (structure: $t = 1.71$, $p < .1$; "speaker": $t = -1.92$, $p < .1$). Readers displayed longer first-fixation durations and gaze durations, but significantly shorter total reading times on the second noun when they encountered SRC sentences. They also had longer total reading times when the sentence was preceded by the non-native speaker's name. "Speaker" marginally affected go-past times ($t = -1.71$, $p < .1$). There was no significant effect on skipping, regressions-in, or regressions-out in this region.

In the main clause region, only "speaker" marginally influenced total reading times ($t = -1.89$, $p < .1$). Readers spent more time processing this region when the sentence had been preceded by the non-native speaker's name regardless of plausibility or structure. No other measures were significant.

Response accuracy data
The best fitted logit model included plausibility ($z = 10.17$, $p < .001$), syntactic structure ($z = 6.67$, $p < .001$), trial order ($z = 2.54$, $p < .05$), and the interaction of plausibility and syntactic structure ($z = -4.08$, $p < .001$) as fixed effects, and random slopes and intercepts for participants and items. Comprehension was more accurate for plausible sentences than for implausible sentences and for SRCs than for ORCs. Items presented later in the experiment were generally comprehended more accurately than items presented earlier in the experiment. Implausible ORC sentences were harder to comprehend than plausible SRC sentences. "Speaker" was not a significant predictor of accuracy.

COMPARISON OF ALL THREE EXPERIMENTS

Experiment 1 examined readers' eye movement patterns and comprehension of the target sentences in normal silent reading. The results largely replicated previous work showing that syntactic structure and plausibility affect reading times and comprehension accuracy. Object-relative clauses were read more slowly than subject-relative clauses (e.g., Gibson, 1998; Traxler et al., 2002). Implausible sentences yielded longer reading times and lower accuracy than plausible ones (Christianson et al., 2010; Ferreira, 2003; Lim & Christianson, 2013b), with implausible ORCs the most likely to be misinterpreted. Experiments 2 and 3 investigated APS effects during silent reading, using different methods to encourage readers to perform APS: photos of "speakers" in Experiment 2; recordings of "speakers" saying their names in Experiment 3. Results from these two experiments indicated that both cues triggered APS effects, such that readers read faster when they perceptually simulated a native speaker's voice while processing the sentences, reflecting the speaking rates of the respective speakers.

One of the main goals of the present study, however, was to determine how APS affected comprehension—specifically, whether inducing a rich prosodic contour would lead to deeper, less good-enough syntactic processing. In order to answer this question, we compared the reading rate and comprehension accuracy data from Experiments 2 and 3 separately to the data from Experiment 1, in which readers were not prompted to engage in APS. In the APS experiments (Experiments 2 and 3), half of the time readers activated the APS of a native speaker's voice, and the other half of the time they simulated the non-native speaker's voice. Thus, we compared the results from trials when participants had been cued to simulate native speech in APS experiments to the same items in the no APS data in Experiment 1, and we then did the same for items in which participants had been cued to simulate non-native speech. Two new datasets were created, and a new variable "experiment" was added to investigate differences across experiment (perceptual simulation of native speech vs. no perceptual simulation; perceptual simulation of non-native speech vs. no perceptual simulation). Linear mixed-effects models were built to analyse reading times, and logit mixed-effects models were fitted to examine accuracy data.

Native "speaker" in Experiment 2 versus no APS in Experiment 1

Sentence reading time

The best fitted model indicated that perceptual simulation (experiment, $t = 2.67$, $p < .001$), structure ($t = -6.20$, $p < .001$), plausibility ($t = -4.57$, $p < .001$), and trial presentation order ($t = -10.28$, $p < .001$) all significantly affected sentence reading time. The critical finding is that participants spent significantly *more* time reading sentences in Experiment 1 when they were *not* perceptually simulating native speech than did participants who were simulating native speech in Experiment 2.

Response accuracy

Results showed that response accuracy when perceptually simulating native speech in Experiment 2 ($M = .87$, $SD = .33$) was significantly higher ($z = -4.9$, $p < .001$) than when not simulating any voice in Experiment 1 ($M = .80$, $SD = .40$).

Non-native "speaker" in Experiment 2 versus no APS in Experiment 1

Sentence reading time

Results for participants who perceptually simulated the non-native speaker's voice in Experiment 2 were compared with the results from Experiment 1. The best fitted model included experiment ($t = 2.68$, $p < .001$), structure ($t = -6.21$, $p < .001$), plausibility ($t = -4.24$, $p < .001$), trial presentation order ($t = -10.90$, $p < .001$), and the interaction between structure and trial presentation order ($t = 4.40$, $p < .001$). Just as with the native speaker APS effects, participants spent significantly *more* time reading sentences when they were *not* perceptually simulating non-native speech.

Response accuracy

Logit mixed model results demonstrated that experiment ($z = -4.07$, $p < .001$), structure ($z = 9.01$, $p < .001$), plausibility ($z = 8.15$, $p < .001$), trial presentation order ($z = 4.03$, $p < .001$), and the interaction of structure and plausibility ($z = -2.10$, $p < .05$) significantly influenced response accuracy across experiments. No APS during silent reading resulted in lower accuracy than APS of non-native speech. In other words, perceptual simulation of both non-native speech and native speech improved readers' comprehension.

Native "speaker" in Experiment 3 versus no APS in Experiment 1

Sentence reading time

The best fitted model demonstrated that plausibility ($t = -7.94$, $p < .001$), structure ($t = -6.40$, $p < .001$), and trial order ($t = -11.26$, $p < .001$) were significant predictors of sentence reading time. Experiment (native speech APS vs. no APS) was not a significant predictor, despite the numerical trend of shorter reading times in the APS condition in Experiment 3 ($M = 4185.53$, $SD = 2245.45$) compared to no APS in Experiment 1 ($M = 4519.61$, $SD = 2435.05$).

Response accuracy

APS ($z = -2.98$, $p < .001$) significantly influenced accuracy across experiments in addition to plausibility ($z = 11.69$, $p < .001$), structure ($z = 12.71$, $p < .001$), trial presentation order ($z = 3.31$, $p < .001$), and the interaction of plausibility and structure ($z = -4.18$, $p < .001$). Accuracy was significantly higher when readers were perceptually simulating the native speaker's voice in Experiment 3 ($M = .88$, $SD = .32$) than for normal silent reading in Experiment 1 ($M = .80$, $SD = .40$).

Non-native "speaker" in Experiment 3 versus no APS in Experiment 1

Sentence reading time

Results showed that plausibility ($t = -7.89$, $p < .001$), structure ($t = -6.82$, $p < .001$), and trial presentation order ($t = -11$, $p < .001$) were significant predictors of sentence reading times. There was no effect of experiment—that is, no significant speed difference between normal silent reading in Experiment 1 ($M = 4519.61$, $SD = 2435.05$) and perceptual simulation of the non-native speaker's voice in Experiment 3 ($M =$

4304.84, $SD = 2152.67$), again despite a numerical trend towards an APS speed advantage.

Response accuracy

Experiment ($z = -2.61$, $p < .001$), sentence structure ($z = 12.04$, $p < .001$), plausibility ($z = 11.09$, $p < .001$), trial presentation order ($z = 3.69$, $p < .001$), and the interaction of structure and plausibility ($z = -2.93$, $p < .001$) significantly affected response accuracy across experiments. The results indicated that APS of non-native speech ($M = .88$, $SD = .33$) was associated with significantly increased response accuracy compared to no perceptual simulation ($M = .80$, $SD = .40$).

APS (EXPERIMENTS 2 AND 3) VERSUS NORMAL SILENT READING (EXPERIMENT 1)

Comparisons of total sentence reading times and response accuracy across Experiments 1 and 2 revealed that when APS was triggered by a speaker's photo, readers tended to read more quickly than under normal silent reading conditions. There are two possible explanations for differences in reading time and accuracy across these experiments. The first explanation, which we prefer, is that APS activated phonological and prosodic information that linked to the syntactic representation and strengthened it, facilitating deeper, less good-enough processing of the sentences and resulting in better comprehension. An alternative explanation is that the cues used to encourage APS drew attention to the reading task, leading to more careful reading and more cautious interpretation. We find this alternative explanation less plausible. Under this explanation, faster reading would need to be characterized as "more careful" or "more cautious", and we are not aware of any accounts of reading in which reading more quickly entails reading more carefully. Triggering APS with either the native or the non-native speakers' name in Experiment 3 resulted in reading that was numerically, but not statistically, also faster than normal silent reading in Experiment 1. Comparisons between Experiment 1 and 3 also demonstrated that perceptual simulation of either native or non-native speech during silent reading led to better comprehension accuracy, which, we contend, is due to deeper sentence processing when readers are engaged in APS.

Comparison of interest areas across experiments

The semantically implausible and syntactically complex sentences used in this study often led to "good-enough" processing (Ferreira et al., 2002), in which readers reached the wrong interpretation based largely on heuristic semantic information while overlooking syntactic structure. To investigate why readers had better comprehension when they activated APS in Experiments 2 and 3, we examined how readers processed these sentences by comparing their eye movement patterns in the four interest areas across the three experiments. Table 2 illustrates the means and standard deviations of each variable in the four regions. Table 3 demonstrates the significant predictors on seven measures in the three experiments. Due to space limitations, we only discuss the eye movement patterns on the second noun region, which was consistent on the whole with the other areas. The detailed modelling results for each interest area are presented in the online Supplemental Material.

In the second noun region, plausibility was the only significant factor that affected all seven measures in Experiment 1. In contrast, structure was the dominant factor that drove eye movements in Experiment 2. The ORC condition yielded longer first-fixation durations, gaze durations, and total reading times on the second noun. It also resulted in more regressions into and out of the region as well as fewer skips. In Experiment 3, both structure and "speaker" modulated first-fixation durations, gaze durations, and total reading times. "Speaker" also significantly affected go-past times on the second noun. It thus appears, again, that in Experiments 2 and 3, readers were more fully committed to the syntactic processing route when they were perceptually simulating speech

Table 3. *Main factors that predicated the first-fixation duration, gaze duration, go-past time, total times, regressions-in, regressions-out, and skipping in the four interest area regions across three experiments*

Expt	*Region*	*FFD*	*GD*	*GP*	*TT*	*RI*	*RO*	*SK*
1	FN				plaus			
	RCV	struc	struc	struc, plaus, Struc × Plaus	struc, plaus, Struc × Plaus	plaus^, struc^	struc^	plaus^, struc
	SN	plaus	plaus	plaus	plaus	plaus	plaus	plaus
	MC		plaus	plaus, struc	plaus, struc			
2	FN		"speaker"		"speaker"	struc		
	RCV	struc	struc	struc	struc	plaus,struc, "speaker"	struc, "speaker"	struc
	SN	struc	struc		struc	struc	struc	struc
	MC	struc			"speaker"			
3	FN		"speaker"		struc			
	RCV	struc	struc	struc	struc, plaus		struc	
	SN	struc, "speaker"	struc, "speaker"	struc, "speaker"	"speaker"			
	MC				"speaker"			

Note: Expt = experiment; FFD = first-fixation duration; GD = gaze duration; GP = go-past time; TT = total times; RI = regressions-in; RO = regressions-out; SK = skipping; FN = first noun; RCV = relative clause verb; SN = second noun; MC = main clause; RC = relative clause; plaus = plausibility; struc = structure.
^$p < .1$ (marginally significant).

during silent reading. In contrast, when not engaged in APS (Experiment 1), readers were more influenced by plausibility heuristics. This effect was not just observed in total sentence reading times and offline measures, but also in individual interest areas and on early measures of processing. This pattern of results suggests that APS was affecting the moment-by-moment processing, not just the offline memory for the input. In other words, the processing in Experiment 1 more closely resembled good-enough processing (Christianson et al., 2001, 2010; Christianson et al., 2006; Ferreira, 2003; Ferreira et al., 2002; Ferreira & Patson, 2007). This difference helps explain why the APS of both native and non-native speech in Experiments 2 and 3 was associated with more accurate comprehension: Readers relied more on the syntactic information, which is the more reliable route to the intended meaning of the sentences. In Experiment 1, on the other hand, readers tended to pursue a good-enough interpretation based on plausibility.

GENERAL DISCUSSION

This study aimed to investigate how auditory perceptual simulation (APS) affects language processing and comprehension during silent reading. We collected online eye movement measures and offline comprehension measures to address the following questions: (a) whether readers would perceptually simulate distinct speakers during silent reading; and (b) what the differences between perceptually simulated speech during silent reading and normal silent reading without perceptual simulation are.

Eye movement data from Experiments 2 and 3 demonstrated that a perceptually simulated "speaker's" voice (which was not physically presented during sentence reading) had significant effects on sentence reading times and response accuracy. Analyses of reading times in the last two experiments demonstrated again and again that participants read sentences that were presented as being spoken by the slower speaker (the non-native speaker) more slowly than those presented as being spoken by the faster speaker (the native speaker). We attribute this effect to APS of the respective speakers' voices during silent reading. In addition, analyses of specific regions also demonstrated that "speaker" had significant effects on earlier and later measures throughout the sentences. Furthermore, split analyses of trials immediately preceding and immediately following "refresher" recordings of native and non-native speaker speech revealed that "speaker" effects weakened over time but then reappeared after exposure to another set of recordings (see online Supplemental Material). These results all support previous observations that APS of speech modulates reading processes (Kosslyn & Matt, 1977; Stites et al., 2013; Yao & Scheepers, 2011).

The most important finding of the current study is that APS facilitated more syntactically driven, less good-enough processing of the sentences, as evidenced by higher response accuracy in the last two experiments under perceptual simulation conditions. By comparing accuracy between perceptually simulated native and non-native speech in the last two experiments with no APS in Experiment 1, we found that readers derived significantly more accurate interpretations of the sentences when they were perceptually simulating either speaker while reading. Further analyses of the interest areas in these three experiments provided a clear explanation for the findings. Eye movement patterns were significantly affected by the syntactic structure in Experiments 2 and 3: ORC sentences produced longer early and late measures along with more regressions into and out of the four interest areas than did SRC sentences. In contrast, in Experiment 1 plausibility was the most consistently significant predictor: When sentences were plausible, readers displayed significantly shorter first-fixation times, gaze durations, go-past times, and total fixation times on the target regions and also tended to make fewer regressions into and out of the regions. Taken together, this constellation of effects suggests that readers tended to rely on "fast and frugal" plausibility heuristics (Ferreira et al., 2002) to interpret the sentences in Experiment 1 rather than the algorithmic syntactic processing stream that was more influential in Experiments 2 and 3. Thus, the results suggest that perceptual simulation of either native or non-native speech facilitated deeper sentence processing and thereby improved comprehension.

One possible explanation of this perceptual simulation advantage is that readers activated richer phonological and prosodic information when they engaged in APS as they read (cf. Breen, 2014). This information, such as the prosodic phrasing of phrasal units, helped solidify the syntactic structure online, and potentially in memory, and reduced the likelihood that the roles of the thematic agents and patients would be reversed. Importantly, the fact that reading speeds were faster in Experiment 2 than in Experiment 1, and no slower in Experiment 3 than in Experiment 1, strongly suggests that cueing APS did not simply cause participants to slow down and "read more carefully". Indeed, the more structurally focused reading in the second two experiments could be considered to be "more careful", but not because more attention was devoted to the text in general. Instead, APS appears to have focused attention on the structural properties of the text, whereas normal silent reading—which certainly also included some degree of typical "inner speech" (Rayner et al., 2012) and "implicit prosody" (Fodor, 2002), but not APS—distributed attention across both structural and, crucially, heuristic lexical–semantic sources of information (cf. Christianson et al., 2010), which are proposed to interfere with one another.

There are two alternative explanations of the results presented here, both of which are predicated on readers' uncertainty about previous material. Neither of these is in principle incompatible with the good-enough account proposed here, but neither can account for the present results without raising questions that cannot be resolved without future research. The first is the "noisy channel" hypothesis (Gibson, Piantadosi, Brink, Bergen, Lim, & Saxe, 2013). Gibson and colleagues (2013) propose that, because meaning is conveyed through "noisy" channels, comprehenders "choose a representation that maximizes meaning recoverability" (p. 1080). In the case of the implausible but reversible relative clauses used here, the hypothesis holds that the more typical thematic role relations between the nouns and the more canonical agent–patient linear order (Townsend & Bever, 2001) should be the ones that readers choose under the noisy conditions of implausibility and noncanonical argument order. This account raises a number of unanswered questions, however, and also requires several assumptions. First, the language attitude surveys (see online Supplemental Material) revealed that participants were very aware of the non-native speaker's accent and generally considered the non-native speaker to be less comprehensible, less reliable, less confident, and less likeable than the native speaker. Presumably, this overall attitude would have introduced more "noise" into the channel. If participants were perceptually simulating both voices, we assume that they would also have been simulating the non-native speaker's accent to some extent, not just the speech rate. (Indeed, the difficulty of simulating the accent may partially explain the overall slower reading rate when simulating the non-native speaker's voice; current experiments—Zhou & Christianson, 2014—are underway to examine the factor of accent vs. speech rate in APS.) The noisy channel hypothesis would predict that reading in the absence of all of this extra "noise"—that is, in Experiment 1—should have been faster and/or more accurate. This was clearly not the case. Perhaps, however, the addition of stronger phonological and prosodic "channels" merged with the syntactic "channel" to produce a stronger "megachannel" and thereby reduce overall noise. We are not aware of research within the noisy channel hypothesis showing that aligning multiple channels reduces noise, and hence uncertainty, but this would certainly be testable.

The second uncertainty-driven account is sketched out by Levy, Bicknell, Slattery, and Rayner (2009), who model how regressive eye movements can be derived from uncertainty about previous syntactic structure. Although Levy et al. did not address semantic uncertainty, the account could be extended to uncertainty of all kinds, similar to the noisy channel hypothesis. Under this view, some of the data reported here might be interpreted as follows: In Experiment 1, uncertainty derives mostly due to implausibility, and also from noncanonical structure (depending on the region one considers). But in Experiments 2 and 3, the uncertainty stems mostly from syntactic structure. Although plausibility affected three measures in Experiment 2, these effects were marginal, and they did not even approach significance in Experiment 3. Thus, under this view, APS served to reduce overall uncertainty and led to better comprehension. Several aspects of this account do not appear consistent with the data, however. First, by bolstering the prosodic representation through APS, uncertainty about plausibility seems to go away, rather than uncertainty about syntax, as syntax is still a main factor in predicting regressions into and out of most regions in Experiments 2 and 3. It is not clear why strengthening the prosodic representation, which is tightly linked to the syntactic representation (Bader, 1998), should not have instead reduced uncertainty about the syntax in the latter two experiments, leaving behind only effects of plausibility. Second, if participants were simply not sure what the identity of the first noun was, or which noun was which, there should have been more regressions into the first noun region in Experiment 1 than in 2 or 3. This was not the case. Also, regressions-out of the final region, which might be predicted if readers were simply unsure what was going on by the time they got to the end of the sentence, were not affected by any of the factors in the three-experiment comparison.

The failure of uncertainty-based approaches to straightforwardly handle multiple aspects of the data here lead us to return to good-enough processing: The more veridical prosodic and phonological representation generated when readers performed APS buttressed the syntactic structure against intrusions from a competing, heuristic-based interpretation. ORCs were still more difficult to process than SRCs (for one or more of the reasons proposed by researchers cited in the introduction), but this computation was no longer significantly affected by the added factor of implausibility. Furthermore, we take the results of the interest area analyses as evidence that this buttressing occurred online, during reading.

CONCLUSION

This study demonstrated that during silent reading readers engaged in auditory perceptual simulation (APS) of the speech of individual "speakers" to whom texts had been attributed. APS effects were observed in both online reading rates and eye movement patterns, and in offline comprehension measures. This perceptual simulation facilitated sentence processing that was tightly linked to the syntactic representation and, consequently, improved comprehension. The results of these studies have clear pedagogical implications and should be of significant interest to educators, linguists, and psychologists who are concerned with how to improve reading comprehensions. Our results show that readers read with better comprehension, and no more slowly, when they activate an auditory simulation of another speaker's voice during silent reading. It is possible that if educators and parents instruct children to actively simulate their teachers' or another fluent English speaker's voice during silent reading, children could practise reading more effectively by themselves both at school and at home. Future research will extend these methods to a variety of readers and texts in order to determine whether APS can be applied to educational settings. If APS is found to consistently improve reading comprehension across a wider variety of texts and readers (including non-native speakers), research focus can shift towards developing pedagogical methods designed to naturally induce students to undertake APS as part of day-to-day silent reading.

Supplemental material

Supplemental material is available via the "Supplemental" tab on the article's online page (http://dx.doi.org/10.1080/13506285.2015.1018282).

REFERENCES

Alexander, J. D., & Nygaard, L. C. (2008). Reading voices and hearing text: Talker specific auditory imagery in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 446–459.

Ashby, J., & Clifton, C. Jr. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, *9*(3), B89–B100.

Ashby, J., Sanders, L. D., & Kingston, J. (2009). Skilled readers begin processing sub-phonemic features by 80 ms during visual word recognition: Evidence from ERPs. *Biological psychology*, *80*(1), 84–94.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baddeley, A., Eldridge, M., & Lewis, V. (1981). The role of subvocalisation in reading. *The Quarterly Journal of Experimental Psychology*, *33*(4), 439–454.

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. *Reanalysis in Sentence Processing Studies in Theoretical Psycholinguistics*, *21*, 1–46.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Barsalou, L. W., & Hale, C. R. (1993). Components of conceptual representation: From feature lists to recursive frames. In I. Van Mechelen, J. Hampton, R. Michalski, & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 97–144). San Diego, CA: Academic Press.

Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain & Language*, *112*, 150–158.

Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass*, *8*(2), 37–50.

Breen, M., & Clifton, C. Jr. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of memory and language*, *64*(2), 153–170.

Callen, V. J., Callois, C., & Forbes, P. A. (1983). Evaluation Reactions to accented English: Ethnicity, Sex Role, and Context. *Journal of Cross-Cultural Psychology*, *14*, 407–426.

Carlson, K., Clifton, C. Jr., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, *45*(1), 58–81.

Christianson, K. (2008). Sensitivity to Syntactic changes in garden path sentences. *Journal of Psycholinguistic Research*, *37*, 391–403.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*, 368–407.

Christianson, K., & Luke, S. G. (2011). Context strengthens initial misinterpretations of text. *Scientific Studies of Reading*, *15*(2), 136–166.

Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects of plausibility on structural priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 538–544.

Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F. (2006). Misinterpretations of garden-path sentences by older and younger adults. *Discourse Processes*, *42*, 205–238.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech*, *40*(2), 141–201.

De Villiers, J. G., Tager, H. B., Flusberg, T., Hakuta, K., & Cohenl, M. (1979). Children's comprehension of relative clauses. *Journal of psycholinguistic Research*, *8*, 499–518.

Drumm, A. M., & Klin, C. M. (2011). When story characters communicate: readers'representations of characters' linguistic exchanges. *Memory & cognition*, *39* (7), 1348–1357.

Edfelt, A. W. (1960). *Silent speech and silent reading*. Chicago: University of Chicago Press.

Edwards, J. R. (1977). Students' reaction to Irish regional accent. *Language & Speech*, *20*, 280–286.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Ferreira, F., & Patson, N. D. (2007). The "good-enough" approach to language comprehension. *Language and Linguistics Compass*, *1*, 71–83.

Filik, R., & Barber, E. (2011) Inner speech during silent reading reflects the reader's regional accent. *PLoS ONE*, *6*(10), e25782. doi:10.1371/journal.pone.0025782

Fodor, J. D. (2002, April). *Psycholinguistics cannot escape prosody*. Paper presented at the meeting of Speech Prosody, Aix-en-Provence, France.

Gass, S., & Varonis, E. (1984). The effect of familiarity on comprehension of the non-native speech. *Language and Speech*, *34*, 65–89.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clause. *Journal of Memory and Language*, *58*, 161–187.

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological science*, *24*(7), 1079–1088.

Gibson, E. A. F. (1998). Linguistic complexity: Locality and syntactic dependencies. *Cognition*, *68*, 1–76.

Giles, H. (1972). Evaluation of personality content from accented speech as a function of listeners' social attitudes. *Perceptual and Motor Skills*, *34*, 168–170. doi:10.2466/pms.1972.34.1.168

Glenberg, A. M., & Kaschak, M. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory & Language*, *43*, 379–401.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97–114.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *32*, 1304–1321.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, *13*, 425–430.

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*, 302–329.

Huey, E. B. (1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press. (Original work published 1908).

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59*, 434–446.

Kim, J. H., & Christianson, K. (2013). Sentence complexity and working memory effects in ambiguity resolution. *Journal of Psycholinguistics Research*, *42*, 393–411.

Kosslyn, S. M., & Matt, M. C. (1977). If you speak slowly, do people read your prose slowly? Person-particular speech recoding during reading. *Bulletin of the Psychonomic Society*, *9*, 250–252.

Kurby, C. A., Magliano, J. P., & Rapp, D. N. (2009). Those voices in your head: Activation of auditory images during reading. *Cognition*, *112*, 457–461.

Levy, B. A. (1975). Vocalization and suppression effects in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(3), 304–316.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, *106*(50), 21086–21090.

Lim, J. H., & Christianson, K. (2013a). Integrating meaning and structure in L1–L2 and L2–L1 translations. *Second Language Research*, *29*, 233–256.

Lim, J. H., & Christianson, K. (2013b). Second language sentence processing in reading for comprehension and translation. *Bilingualism: Language and Cognition*, *16*, 518–537.

Louwerse, M. M., & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science*, *35*, 381–398.

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, *114*, 96–104.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*, 477–494.

Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F. (2009). Lingering misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 280–285.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, *14*, 119–124.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, *11*, 164–167.

Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. New York: Cambridge University Press.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. E. Jr. (2012). *The psychology of reading*. New York: Psychology Press. ISBN: 978-1-84872-975-9.

Sachs, J. S. (1967). Recopition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, *2*(9), 437–442.

Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, *69*(2), 104–120.

Slowiaczek, M. L., & Clifton, C. Jr. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior*, *19*(5), 573–582.

Solomon, K. O., & Barsalou, L. W. (2004). Perceptual simulation in property verification. *Memory & Cognition, 2004*, *32*, 244–259.

Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, *25*(2), 249–271.

Stites, M., Luke, S. G., & Christianson, K. (2013). The psychologist said quickly, "dialogue descriptions modulate reading speed!" *Memory and Cognition*, *41*, 137–151.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*(1), 201–216.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules.* Cambridge, MA: MIT Press.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*, 69–90.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory & Language*, *55*, 157–166.

van Gompel, R. P., Pickering, M. J., Pearson, J., & Jacob, G. (2006). The activation of inappropriate analyses in garden-path sentences: Evidence from structural priming. *Journal of Memory and Language*, *55*(3), 335–362.

Vornik, L., Sharman, S., & Garry, M. (2003). The power of the spoken word: Sociolinguistics cues influence the misinformation effect. *Memory*, *11*, 101–109.

Watson, D., & Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica*, *59*(2–3), 279–300.

Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology*, *36*, 559–570.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*, 250–271.

Yao, B., & Scheepers, C. (2011). Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition*, *121*, 447–453.

Zhou, P., & Christianson, K. (2014, September). *Mechanisms underlying auditory perceptual simulation: reading speed or speech biases?* Poster presented at Architectures and Mechanisms for Language Processing. Edinburg, UK.

Zhou, P., & Christianson, K. (2015, in preparation). Comparison of attitudes towards native and non-native English speech.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 44, pp. 35–62). San Diego, CA: Academic.

Zwaan, R. A. (2008). Experiential traces and mental simulations in language comprehension. In M. DeVega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 165–180). Oxford, UK: Oxford University Press.

Zwaan, R. A. (2009). Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology*, 39(7), 1142–1150.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

APPENDIX

Stimuli for Experiments 1, 2, and 3

Plausible and subject-relative clause sentences:

1. The police officer that arrested the citizen was handsome.
2. The chef that ruined the food was in the kitchen.
3. The cop that pursued the thief was driving a car.
4. The cat that chased the mouse was fast.
5. The ghost that scared the boy was hiding behind a curtain.
6. The bird that ate the worm was small.
7. The hunter that shot the deer was in the Rocky Mountains.
8. The lawyer that sued the doctor was smart.
9. The coach that scolded the player won the championship twice.
10. The owner that fed the cat was sitting on a sofa.
11. The man that walked the dog was in the park.
12. The father that scolded the teenager was in the living room.
13. The guide that led the tourist liked Europe a lot.
14. The waiter that served the guest was tall.
15. The soldier that protected the villager was brave.
16. The fly that ate the frog was green.
17. The doctor that treated the patient was female.
18. The teacher that quizzed the student was in the classroom.
19. The angler that caught the fish was in the middle of the ocean.
20. The detective that investigated the suspect was very tired.
21. The dog that bit the man was in the yard.
22. The golfer that hit the ball was in the shade.
23. The politician that the voter deceived was Korean.
24. The mother that bathed the child smelled nice.
25. The bird that protected the chick was in the big tree.
26. The reporter that interviewed the actress was at the coffee shop.
27. The fan that admired the Hollywood star loved to wear big sunglasses.
28. The dog that herded the sheep was very furry.
29. The grandmother that dressed the child had a beautiful smile.
30. The tutor that taught the student solved the math problem.
31. The consultant that advised the client was very clever.
32. The homeowner that the gardener paid loved the garden.
33. The kids that obeyed the teacher enjoyed the summer break.
34. The criminal that kidnapped the girl was on CNN news.
35. The boss that fired the worker was unhappy.
36. The volunteer that helped the blind was very handsome.
37. The fan that cheered for the baseball player wore a red shirt.
38. The parent that raised the twins lived in New York.
39. The parents that punished the child went to church.
40. The dean that awarded the student a prize was very famous in the school.
41. The nuns that took care of the orphans live in a small village.
42. The lawyer that defended the client was worried about the result.
43. The secretary that assisted the President was in a conference room.
44. The boy that petted the puppy had cute eyes.
45. The terrorist that held the hostage was located in the building.
46. The wrecker that towed the car drove at a high speed.
47. The conductor that led the orchestra was pleased with the performance.
48. The guard that locked up the prisoner regretted past decisions.