

Interactions between distal speech rate, linguistic knowledge, and speech environment

Tuuli Morrill¹ · Melissa Baese-Berk² · Christopher Heffner³ · Laura Dilley⁴

Published online: 21 March 2015
© Psychonomic Society, Inc. 2015

Abstract During lexical access, listeners use both signal-based and knowledge-based cues, and information from the linguistic context can affect the perception of acoustic speech information. Recent findings suggest that the various cues used in lexical access are implemented with flexibility and may be affected by information from the larger speech context. We conducted 2 experiments to examine effects of a signal-based cue (distal speech rate) and a knowledge-based cue (linguistic structure) on lexical perception. In Experiment 1, we manipulated distal speech rate in utterances where an acoustically ambiguous critical word was either obligatory for the utterance to be syntactically well formed (e.g., Conner knew that bread and butter (are) both in the pantry) or optional (e.g., Don must see the harbor (or) boats). In Experiment 2, we examined identical target utterances as in Experiment 1 but changed the distribution of linguistic structures in the fillers. The results of the 2 experiments demonstrate that speech rate and linguistic knowledge about critical word obligatoriness can both influence speech perception. In addition, it is possible to alter the strength of a signal-based cue by changing information in the speech environment. These results provide

support for models of word segmentation that include flexible weighting of signal-based and knowledge-based cues.

Keywords Speech perception · Spoken word recognition · Speech rate

Listeners are constantly encountering variability in the speech signal; the acoustic environment, including both speech and nonspeech sounds, is always in flux, and unlimited combinations of words can be produced by an interlocutor. Lexical access is thus a complicated process since units of lexical meaning must be extracted from a continuous acoustic signal. Listeners use multiple sources of information to understand speech. Acoustic information, including spectral and temporal characteristics of the speech signal, is generally referred to as “signal-based” information. On the other hand, “knowledge-based” information used in speech processing and segmentation can include both phonological structure and phonotactics (McQueen, 1998), as well as syntactic structure, semantic context, and other pragmatic factors (Marslen-Wilson & Welsh, 1978). Investigations of how listeners integrate signal-based and knowledge-based cues in lexical access have shown that the effects of linguistic context on the perception of acoustic information are quite powerful. Lexical access is facilitated by syntactic and semantic predictability (Staub & Clifton, 2006), and listeners can “repair” missing phonological material when there are other cues to the lexical content of an utterance (Warren & Sherman, 1974; Samuel, 2001).

A body of work has highlighted the need for examination of the relationship between signal-based and knowledge-based information, as well as speech environment (or experimental context), in lexical processing. Models of word segmentation have generally proposed hierarchically organized processing of the different sources of information affecting

✉ Tuuli Morrill
tmorrill@gmu.edu

¹ Program in Linguistics, Department of English, George Mason University, 4400 University Drive, Fairfax, VA 22030, USA

² Department of Linguistics, 1290 University of Oregon, Eugene, OR 97403, USA

³ Program in Neuroscience and Cognitive Science, University of Maryland, 1401 Marie Mount Hall, College Park, MD 20742, USA

⁴ Department of Communicative Sciences and Disorders, Michigan State University, Oyer Center, 1026 Red Cedar Road, East Lansing, MI 48824, USA

lexical access (Gow & Gordon, 1995; Mattys, White, & Melhorn, 2005; Norris, McQueen, Cutler, & Butterfield, 1997). For example, Mattys and colleagues (2005) proposed a model in which the information integrated in lexical processing was hierarchically organized into “tiers,” with knowledge-based factors such as structural (syntactic and semantic) and lexical knowledge comprising the most important sources of information. However, other lines of research have demonstrated that acoustic information strongly influences lexical recognition, especially for reduced word forms in casual speech (e.g., Ernestus, Baayen, & Schreuder, 2002); in fact, when syntactic and lexical cues conflicted with acoustic information, recognition was affected more by acoustic cues than structural cues (Van De Ven, Ernestus, & Schreuder, 2012). These findings are in line with Mattys and colleagues’ postulation in a revision of their model that a strictly hierarchical organization cannot account for changes in sensitivity to both acoustic and syntactic information in the speech context (Mattys & Melhorn, 2007; Mattys, Melhorn, & White, 2007). Mattys and colleagues found “compensatory segmentation strategies” in listeners’ perception of lexically ambiguous sequences (e.g., *plum#pie* or *plump#eye*). Generally, lexical (semantic) information decreased listeners’ sensitivity to acoustic information (i.e., they relied on the available lexical information); however, if acoustic cues were strong, effects of the sentential or lexical context were smaller. Mattys et al. noted that these data suggested graded trading of information between the previously proposed tiers of factors in lexical recognition.

Recently, several lines of research have demonstrated that nonlocal information, specifically the signal-based factors of distal (i.e., nonadjacent context) speech rate and rhythm, can influence spoken word segmentation and recognition (e.g., Baese-Berk et al., 2014; Dilley & Pitt, 2010; Holt, 2005; Morrill, Dilley, McAuley, & Pitt, 2014; Reinisch, Jesse, & McQueen, 2011). For example, the perception of entire syllables or words can be affected by distal speech rate. In a phrase such as “see the harbor or boats,” the monosyllabic “or” can be heavily coarticulated with the surrounding syllables and interpreted as *harbor boats*; a slowed speech rate in the distal context causes listeners to less often report hearing the function word “or” (e.g., Dilley & Pitt, 2010). Several studies have now shown these effects to be robust across contexts and stimulus types. In general, when an ambiguous lexical sequence is located in a region in which the speech rate is *relatively fast* compared to the speech rate in the distal context, the phonemic material in the function word is less likely to be perceived than if the target region occurs at the same rate as the distal context.

Though most studies of distal speech rate and rhythm have examined these cues in relative isolation from knowledge-based cues, there has been one preliminary test of the interaction of knowledge-based and distal signal-based cues.

Manipulation of distal prosodic pitch and duration patterns can affect the segmentation of syllable sequences with ambiguous parses (e.g., four syllables parsed as either *timer, derby or tie, murder, bee*; e.g., Dilley, Mattys, & Vinke, 2010). Explicitly contrasting the effects of semantic context with the effects of distal prosody, Dilley and colleagues found strong effects of distal prosody even when an alternative parse was favored by the semantic context. The finding supports the idea that distal context-based acoustic cues may “trade off” or interact flexibly with linguistic knowledge-based cues, as suggested in the proposal of Mattys and colleagues (Mattys & Melhorn, 2007; Mattys et al., 2007) that listeners use compensatory segmentation strategies (e.g., employing knowledge-based cues to a lesser degree when signal-based cues are strong) which cannot be accounted for by a strictly hierarchical model.

Here, we conducted two experiments to examine the ways in which a distal signal-based cue and linguistic knowledge may interact to influence lexical recognition and ask whether listeners’ use of these cues is influenced by the experimental speech environment. We examine distal rate effects in two types of utterances. The first included a short, coarticulated, acoustically ambiguous critical word that is obligatory for the utterance to be syntactically well-formed (an “obligatory” item, e.g., “Conner knew that bread and butter [are] both in the pantry,” which is not grammatical without the word *are*). The second type of utterance had a critical word that is optional (an “optional” item, e.g., “Don must see the harbor [or] boats,” which is grammatical with or without *or*). Distal speech rate was manipulated in both types of utterances. Consistent with previous findings, we predict an overall lowering of critical word reports in utterances with slowed distal speech rates; if linguistic knowledge about the obligatoriness of the critical word influences listeners’ reports, we predict higher critical word report rates for obligatory items. In the second experiment, we examine whether speech environment can affect the roles that linguistic knowledge and distal speech rate play in lexical access. We examined identical target utterances as in Experiment 1, but changed the distribution of linguistic structures by including filler items that were overtly missing “obligatory” critical words (e.g., “Lily decided to put __ patch on the jacket...”).

Experiment 1

Method

Participants Forty-five participants were recruited for research credit at Michigan State University and the University of Oregon. Participants were continuously recruited during predetermined periods of time during the school semester. All were native speakers of English who self-reported normal hearing and were at least 18 years old.

Stimuli and design A 2 (critical word obligatoriness: optional or obligatory) \times 2 (distal rate: slowed or unaltered) within-subjects design was used. The items in both experiments were drawn from materials collected by Dilley and Pitt (2010). Obligatory items consisted of an utterance in which a critical function word was obligatory for the utterance to be syntactically well-formed (e.g., “Conner knew that bread and butter [are] both in the pantry”). Optional items consisted of an utterance which was syntactically well-formed with or without the critical word (e.g., “Don must see the harbor [or] boats). Items across the Obligatory and Optional conditions were matched for equal numbers of each critical word (*are*, *our*, *or*, *her* and *a*).

Following Dilley and Pitt (2010), utterances were presented at two distal rates: unaltered or slowed. All stimuli included a target region consisting of the critical word and the preceding syllable. For example, in the utterance “Don must see the harbor (or) boats,” the *target* region consisted of the critical word *or*, the preceding syllable *-bor*, and the following phoneme [b] – in this case, [b]. The *context* region included the non-target (preceding and following) portions of the utterance. For slowed rates, the context portions were multiplied by 1.75 using the PSOLA algorithm in Praat (Boersma & Weenink, 2012) so that the duration of the context was 175 % the duration of the original context, thus slowing the speech rate. The unaltered items were multiplied by 1.0 using the PSOLA algorithm, maintaining their original speech rate. Presentation of unaltered or slowed rate for experimental items was counterbalanced across participants and items. Filler items had no intended ambiguity regarding the number of words they contained. They were presented at either the unaltered or the slowed rate; the duration of the entire utterance was manipulated.

Procedure Participants listened to utterances over headphones and typed what they heard (of the entire utterance). The order of trials was pseudorandomized. Participants heard 26 target utterances and 36 filler utterances.

Results and discussion

We analyzed whether participants reported hearing the critical function word (FW) in the acoustically ambiguous region for each utterance. “FW present” was coded for responses in which the produced FW (or another grammatical FW in its place) was reported. “FW absent” was coded for responses in which a FW was not reported. Responses that did not accurately report the preceding and/or following lexical items (the words surrounding the critical FW), were not included in the analysis (10.7 % of responses). Responses were analyzed using mixed-effects logistic regression (see Table 1). The model predicted the likelihood of a critical word report;

Table 1 Means (standard deviations) of proportion of critical word reports in Experiment 1

| | <i>M</i> | <i>SD</i> |
|----------------------|----------|-----------|
| Overall unaltered | .89 | .31 |
| Overall slowed | .41 | .49 |
| Obligatory unaltered | .93 | .26 |
| Obligatory slowed | .54 | .49 |
| Optional unaltered | .85 | .36 |
| Optional slowed | .27 | .44 |

critical word obligatoriness, speech rate, and the interaction between obligatoriness and rate were fixed factors, and included random intercepts for subjects and items.¹

Figure 1 shows the critical word report rate as a function of speech rate (slowed and unaltered) and critical word obligatoriness (obligatory and optional; see Table 1 for descriptive statistics). Examining the descriptive statistics, it is clear that slowed speech rate results in a lower critical word report rate than the unaltered speech rate. The optional sentences appear to have slightly lower critical word report rates than the obligatory sentences.

The results of the logistic regression model reflect these observations (see Table 2). Speech rate was a valid predictor of critical word report rates ($p < .001$). Despite an apparent trend, obligatoriness was not a significant predictor ($p = .252$), nor was the interaction between obligatoriness and speech rate ($p = .117$).

These results demonstrate that the speech rate effect is robust but that critical word obligatoriness was not a significant predictor of report rates. These results raise two main questions. First, was the strength of the signal-based effect of speaking rate due to an experimental context allowing participants to complete the task while ignoring linguistic information? Second, can listeners change the weighting of the signal-based and knowledge-based cues as a function of the speech environment (in the context of an experiment)? To address these questions, we conducted a second experiment in which we altered the types of structures to which listeners were exposed and we changed the task demands.

Experiment 2

In Experiment 2, we utilized the same target utterances as in Experiment 1, changing only the filler items, to examine whether listeners would perceive the stimuli differently when the experimental speech environment changed in the

¹ Models with random slopes for within-subjects factors for both Experiments 1 and 2 exhibited the same pattern of results as without slopes but exhibited worse model fit; therefore, models with the better fit (with only random intercepts) are presented for all analyses.

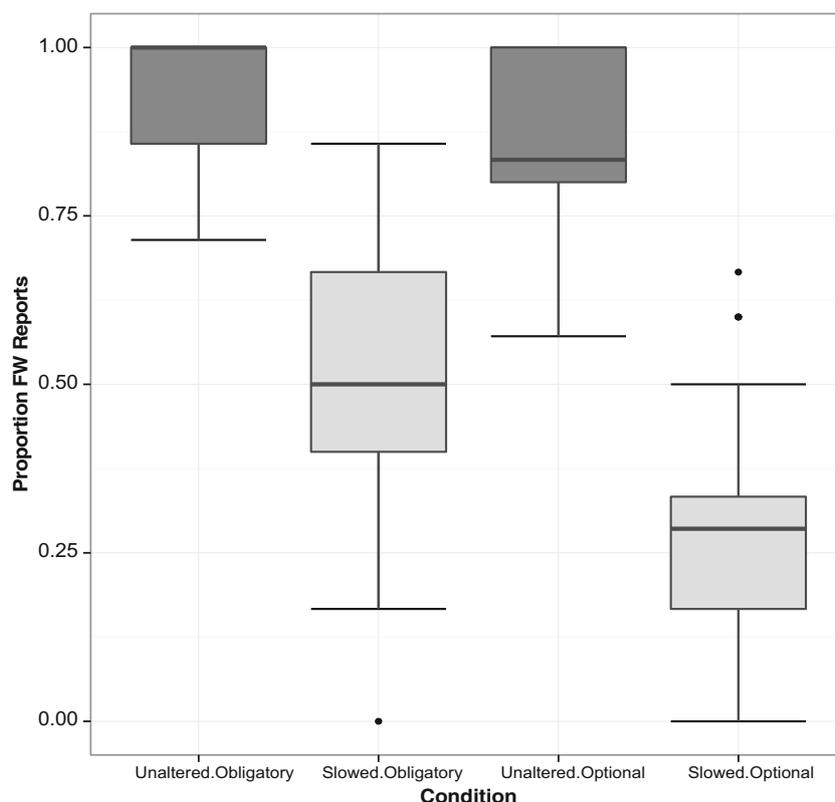


Fig. 1 Proportion of critical function word (FW) reports for each Rate and Obligatoriness condition in Experiment 1

following manner: (1) fillers included utterances which were overtly missing function words and (2) participants were explicitly made aware of the linguistic structure of the stimuli, with a naturalness rating task. We predicted that listeners would exhibit sensitivity to the speech environment and change the weighting of available cues in the following ways: (1) The presence of utterances with overtly missing words, combined with the rating task, would increase listeners' awareness of linguistic structure and result in an effect of syntactic obligatoriness while at the same time; (2) listeners would still exhibit a high degree of sensitivity to signal-based rate information in the environment in which syntactic structure is perceived as unreliable.

Table 2 Estimates of predictor variables and their reliability in the mixed-models analysis for Experiment 1

| | Estimate | SE | z value | p value |
|--------------------------|----------|--------|---------|---------|
| Intercept | 3.6785 | 0.6319 | 5.821 | < .001 |
| Rate – Slowed | -3.5022 | 0.3743 | -9.356 | < .001 |
| Syntactic obligatoriness | -0.9919 | 0.8663 | -1.145 | .252 |
| Rate * Obligatoriness | | | | |
| – Slowed * Optional | -0.8204 | 0.5238 | -1.566 | .117 |

Method

Participants Forty-three participants were recruited for research credit at Michigan State University during a predetermined time period during the school semester. All were native speakers of English who self-reported normal hearing and were at least 18 years old.

Stimuli and design The target stimuli and design were identical to those of Experiment 1. Unlike in Experiment 1, one-third of the filler utterances contained overtly missing critical words, resulting in violations of syntactic well-formedness (e.g., “Lily decided to put ___ patch on the jacket...”). These fillers were similar in structure to the “obligatory” target items, in which the critical word would be necessary to perceive the utterance as well-formed; however, in the fillers, the critical word was not produced and therefore was not present in the acoustic signal.

Procedure The procedure was similar to Experiment 1. Listeners transcribed each utterance. However, before transcribing, they were asked to complete a grammaticality judgment task, rating the naturalness of each utterance on a scale of 1 (*completely unnatural*) to 6 (*completely natural*; see Appendix A). Listeners then heard the utterance a second time and transcribed it.

Results and discussion

Before analyzing the critical word report rates, we examined the naturalness ratings. When listeners did not report a critical function word in syntactically obligatory conditions, they rated those sentences as less grammatical than when they did report a critical word in the obligatory conditions (see Table 3). This suggests that we achieved our goal of having listeners attend to the grammaticality of the items.

As in Experiment 1, we analyzed whether participants reported hearing the critical function word in the acoustically ambiguous region for each utterance (see Table 4 for descriptive statistics). Responses were again analyzed using mixed-effects logistic regression (see Table 5), with identical model structure to Experiment 1.

Figure 2 shows the critical word report rates for target utterances as a function of speech rate and critical word obligatoriness. As in Experiment 1, critical word report rates are lower for the slowed than for the unaltered speech rates. Furthermore, optional sentences have lower critical word report rates than obligatory sentences.

The results of the regression model again support these observations. Similar to Experiment 1, speech rate was a significant predictor of critical word report rates ($p < .001$). Unlike Experiment 1, obligatoriness was also a significant predictor of critical word report rates ($p < .05$). The interaction between speech rate and critical word obligatoriness was not significant ($p = .765$).

The results of Experiment 2 confirm the robust effect of speech rate. In addition, an effect of critical word obligatoriness emerges in this environment. It appears that the critical word report rates are lower for Experiment 2 than Experiment 1, particularly for slowed speech rates. We performed an analysis directly comparing both experiments (see Table 6), with experiment as a fixed factor and interactions between speech rate, critical word obligatoriness, and experiment as additional factors. A significant interaction emerged between experiment and speech rate, indicating that critical word report rates were significantly lower in Experiment 2 than in Experiment 1 in the slowed condition ($p < .01$).

Table 3 Grammaticality ratings in Experiment 2. For target stimuli, the interaction between obligatoriness and critical word report contributes significantly to model fit ($\chi^2 = 22.546$, $p < .001$) in a linear regression analysis; targets are rated higher when listeners report the critical word and the critical word is obligatory

| Condition | Critical word report | Filler $M(SD)$ | Target stimulus $M(SD)$ |
|------------|----------------------|----------------|-------------------------|
| Obligatory | FW No | 3.48 (1.55) | 3.30 (1.48) |
| | FW Yes | 5.22 (1.09) | 4.88 (1.37) |
| Optional | FW No | 4.97 (1.25) | 4.01 (1.66) |
| | FW Yes | 5.02 (1.31) | 4.52 (1.47) |

Table 4 Means (standard deviations) of proportion of critical word reports in Experiment 2

| | M | SD |
|----------------------|-----|------|
| Overall unaltered | .87 | .33 |
| Overall slowed | .23 | .42 |
| Obligatory unaltered | .94 | .24 |
| Obligatory slowed | .29 | .45 |
| Optional unaltered | .79 | .41 |
| Optional slowed | .17 | .38 |

General discussion and conclusion

In the present studies, we examine how the strengths of a linguistic knowledge-based cue (critical word obligatoriness) and a signal-based cue from the nonlocal context (distal speech rate) change across speech environments. In Experiment 1, there was no effect of critical word obligatoriness, but this effect emerges in Experiment 2. Furthermore, comparing the two experiments in which identical target items were used, we see that experimental context may influence the strength of the distal speech rate effect.

These results are in accord with studies showing that listeners adapt to their linguistic environment, changing their expectations and weighting available cues depending on the speech context. Cue-weighting, the use of multiple cues to varying degrees depending on a number of factors, has been utilized to account for a variety of linguistic processing phenomena, ranging from phoneme identification (e.g., Pisoni & Luce, 1987; Repp, 1982) to syntactic structure (e.g., Beach, 1991). Early work by Miller and colleagues showed that while speaking rate and semantic context influence the perception of acoustically ambiguous phonemes (Miller, Green, & Schermer, 1984), the effect of semantic context on VOT categorization was found only when the experimental task included a judgment identifying the semantic context. This suggests that the experimental context in which a task is performed can also influence cue-weighting. Recently, Jaeger and colleagues (Fine & Jaeger, 2013) have shown that over the course of a single experiment listeners can adapt to the frequency of certain syntactic structures in the linguistic environment, resulting in changes in syntactic processing. Additionally, in studies of lexical recognition, participants' speed and accuracy are affected by the type and number of real or nonwords in the experimental context; these findings have been accounted for in models of speech perception which incorporate attentional modulation (e.g., Mirman, McClelland, Holt, & Magnuson, 2008; Pitt & Szostak, 2012). Attentional modulation is considered evidence for interactive processing, resulting from effects of attention at the “lexical layer,” with suggested parallels for larger contextual domains. A potentially related phenomenon has been observed with nonnative

Table 5 Estimates of predictor variables and their reliability in the mixed models analysis for Experiment 2

| | Estimate | SE | z value | p value |
|-----------------------------------|----------|--------|---------|---------|
| Intercept | 3.2547 | 0.4869 | 6.684 | < .001 |
| Rate – <i>Slowed</i> | -4.3670 | 0.3389 | -12.887 | < .001 |
| Syntactic Obligatoriness | | | | |
| – <i>Optional</i> | -1.2918 | 0.6577 | -1.964 | .0495 |
| Rate * Obligatoriness | | | | |
| – <i>Slowed</i> * <i>Optional</i> | 0.1454 | 0.4871 | 0.299 | .7653 |

speakers, who rely less on lexical information than do native speakers when parsing ambiguous speech under a cognitive load – less reliable linguistic knowledge may contribute to this effect (Mattys, Carroll, Li, & Chan, 2010).

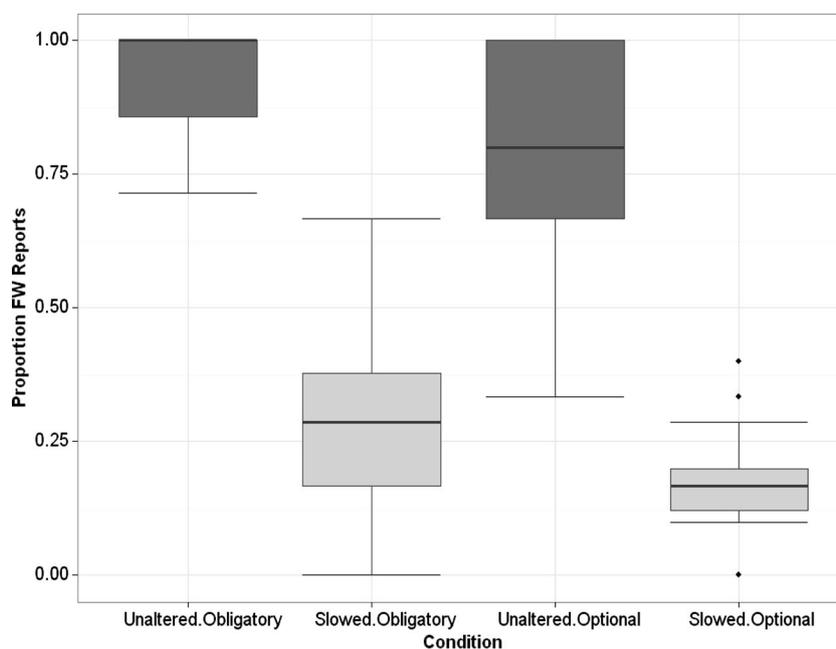
The current results support a growing body of work suggesting that some mechanism for flexibility of cue-weighting, including cues from the speech context, must be incorporated into models of lexical recognition (e.g., Mattys & Melhorn, 2007; Mattys et al. 2007). Differences in perception of identical stimuli across Experiments 1 and 2 suggest that listeners may “weight” the signal-based distal speech rate cue more heavily (regardless of syntactic obligatoriness) when knowledge-based cues are overall less reliable. Recent neuroimaging studies suggest that both knowledge-based information (syntax, semantics) and rhythmical information contribute to listeners’ expectancies about upcoming linguistic input, and that these cues interact at various stages of processing (e.g., Rothermich & Kotz, 2013; Schmidt-Kassow & Kotz, 2009).

We adopt the proposal of recent accounts of distal rate effects, to suggest that neural entrainment, by regularities in

Table 6 Estimates of predictor variables and their reliability in the mixed models analysis comparing Experiments 1 and 2

| | Estimate | SE | z value | p value |
|---|----------|--------|---------|---------|
| Intercept | 3.3156 | 0.5777 | 5.739 | < .001 |
| Rate – <i>Slowed</i> | -3.1917 | 0.4566 | -6.991 | < .001 |
| Syntactic Obligatoriness | | | | |
| – <i>Optional</i> | -0.8902 | 0.7716 | -1.154 | .2486 |
| Experiment – <i>Experiment 2</i> | 0.6166 | 0.6226 | 0.990 | .3220 |
| Rate * Experiment | | | | |
| – <i>Slowed</i> * <i>Experiment 2</i> | -1.7826 | 0.6845 | -2.604 | < .01 |
| Rate * Obligatoriness | | | | |
| – <i>Slowed</i> * <i>Optional</i> | -0.7847 | 0.6348 | -1.236 | .2164 |
| Obligatoriness * Experiment | | | | |
| – <i>Optional</i> * <i>Experiment 2</i> | -0.8699 | 0.7407 | -1.174 | .2402 |
| Rate * Obligatoriness * Experiment | | | | |
| – <i>Slowed</i> * <i>Optional</i> * <i>Experiment 2</i> | 1.1394 | 0.8937 | 1.275 | .2023 |

the speech signal, leads to altered perceptions of lexical information (e.g., Baese-Berk et al., 2014; Dilley & Pitt, 2010; Peelle & Davis, 2012). Listeners’ expectations about the amount of upcoming speech material (and thus the number of syllables and/or words present) are affected by the preceding speech rate; in the current data, these expectations often supersede those based on linguistic structure. It is possible that in environments of less reliable linguistic information, listeners’ reports more often reflect their altered perceptions of the acoustic signal. Thus, these results provide support for models of word segmentation and lexical recognition that

**Fig. 2** Proportion of critical function word (FW) reports for each Rate and Obligatoriness condition in Experiment 2

employ cue-weighting at multiple levels of representation, including acoustic phonetic, semantic, syntactic, and distal information, such as speech rate.

Acknowledgments We would like to thank Amanda DePelsMaeker, Marisa Reynolds, Kathryn Ver Plank, Kellie Voss, Elizabeth Wieland, and Paul Olejarczuk for assistance with this project. This work was partially supported by NSF CAREER award BCS 0874653 to Laura C. Dilley.

Appendix A. Instructions for naturalness rating task

1. Some of these sentences will sound completely natural and well formed, and some of them will sound unnatural or ungrammatical.
2. In addition, some of the sentences are slower than others.
3. You will have two tasks. The first is to rate the naturalness of the sentence on a scale from 1 to 6, where 1 is *completely unnatural* and 6 is *completely natural*.
4. In this task, you should not think about the speed or sound quality of the sentence, only the grammaticality.

References

- Baese-Berk, M. M., Heffner, C. C., Dilley, L. C., Pitt, M. A., Morrill, T. H., & McAuley, J. D. (2014). Long-term temporal tracking of speech rate affects spoken-word recognition. *Psychological Science*, *5*(8), 1546–1553.
- Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, *30*, 644–663.
- Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer* [Computer program]. Software and manual available online at <http://www.praat.org>
- Dilley, L. C., Mattys, S., & Vinke, L. (2010). Potent prosody: Comparing the effects of distal prosody, proximal prosody, and semantic context on word segmentation. *Journal of Memory and Language*, *63*, 274–294.
- Dilley, L. C., & Pitt, M. (2010). Altering context speech rate can cause words to appear or disappear. *Psychological Science*, *21*, 1664–1670.
- Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language*, *81*, 162–173.
- Fine, A. B., & Jaeger, F. T. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*, 578–591.
- Gow, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 344–359.
- Holt, L. L. (2005). Temporally nonadjacent nonlinguistic sounds affect speech categorization. *Psychological Science*, *16*(4), 305–312.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions during word recognition in continuous speech. *Cognition*, *10*, 487–509.
- Mattys, S. L., Carroll, L. M., Li, C. K. W., & Chan, S. L. Y. (2010). Effects of energetic and informational masking on speech segmentation by native and non-native speakers. *Speech Communication*, *52*(11/12), 887–899.
- Mattys, S. L., & Melhorn, J. F. (2007). Sentential, lexical, and acoustic effects on the perception of word boundaries. *Journal of the Acoustical Society of America*, *122*, 554–567.
- Mattys, S. L., Melhorn, J. F., & White, L. (2007). Effects of syntactic expectations on speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 960–977.
- Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, *134*, 477–500.
- McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, *39*, 21–46.
- Miller, J., Green, K., & Schermer, T. M. (1984). A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception and Psychophysics*, *36*(4), 329–337.
- Mirman, D., McClelland, J. L., Holt, L. L., & Magnuson, J. S. (2008). Effects of attention on the strength of lexical influences on speech perception: Behavioral experiments and computational mechanisms. *Cognitive Science*, *32*, 398–417.
- Morrill, T., Dilley, L., McAuley, J. D., & Pitt, M. (2014). Distal rhythm influences whether or not listeners hear a word in continuous speech: Support for a perceptual grouping hypothesis. *Cognition*, *131*, 69–74.
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, *34*, 191–243.
- Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*(320), 1–17. doi:10.3389/fpsyg.2012.00320
- Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition*, *25*, 21–52.
- Pitt, M. A., & Szostak, C. M. (2012). A lexically biased attentional set compensates for variable speech quality caused by pronunciation variation. *Language and Cognitive Processes*, *27*(7/8), 1225–1239. doi:10.1080/01690965.2011.619370
- Reinisch, E., Jesse, A., & McQueen, J. M. (2011). Speaking rate from proximal and distal context is used during word segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 978–996.
- Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, *92*, 81–110.
- Rothermich, K., & Kotz, S. A. (2013). Predictions in speech comprehension: fMRI evidence on the meter-semantic interface. *NeuroImage*, *70*, 89–110.
- Samuel, A. G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science*, *12*, 348–351.
- Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, *21*(9), 1693–1708.
- Staub, A., & Clifton, C. (2006). Syntactic prediction and language comprehension: Evidence from *either . . . or*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 425–436.
- Van De Ven, M., Ernestus, M., & Schreuder, R. (2012). Predicting acoustically reduced words in spontaneous speech: The role of semantic/syntactic and acoustic cues in context. *Labphon*, *3*(2), 455–481. doi:10.1515/lp-2012-0020
- Warren, R. M., & Sherman, G. L. (1974). Phonemic restorations based on subsequent context. *Perception & Psychophysics*, *16*, 150–156.