Moving to the Speed of Sound: Context Modulation of the Effect of Acoustic Properties of Speech

Hadas Shintel, Howard C. Nusbaum

Department of Psychology and Center for Cognitive and Social Neuroscience, University of Chicago

Abstract

Suprasegmental acoustic patterns in speech can convey meaningful information and affect listeners' interpretation in various ways, including through systematic analog mapping of message-relevant information onto prosody. We examined whether the effect of analog acoustic variation is governed by the acoustic properties themselves. For example, fast speech may always prime the concept of speed or a faster response. Alternatively, the effect may be modulated by the context-dependent interpretation of those properties; the effect of rate may depend on how listeners construe its meaning in the immediate linguistic or communicative context. In two experiments, participants read short scenarios that implied, or did not imply, urgency. Scenarios were followed by recorded instructions, spoken at varying rates. The results show that speech rate had an effect on listeners' response speed; however, this effect was modulated by discourse context. Speech rate affected response speed following contexts that emphasized speed, but not without such contextual information.

Keywords: Prosody; Spoken language comprehension; Perception–behavior effect; Discourse context; Embodied cognition

1. Introduction

Beyond words and sentence structure there is a substantial amount of communicative information available in spoken language, thereby setting it apart from written language. For example, prosody affects comprehension by providing a wide range of information, from cueing intended syntactic structure(Beach, 1991; Carlson, Clifton, & Frazier, 2001; Snedeker & Trueswell, 2003; Watson & Gibson, 2004) and discourse structure (Birch & Clifton, 1995; Bock & Mazzella, 1983; Terken & Nooteboom, 1987), to signalling speakers' emotions (Banse & Scherer 1996). Even spontaneous speech disfluencies, hesitations, and pauses are not simply

---

Correspondence should be sent to Hadas Shintel, Center for Cognitive and Social Neuroscience, University of Chicago, Beecher 102, 5848 South University Avenue, Chicago, IL 60637. E-mail: hadas@uchicago.edu

random production flaws, but provide information to listeners and can affect comprehension (Arnold, Tanenhaus, Altman, & Fagnano, 2004; Bailey & Ferreira, 2003; Barr, 2003; Clark & Fox Tree, 2002).

Recently, we found evidence that prosody can also convey semantic-referential information about external objects and events. Speakers can capitalize on cross-modal associations (e.g., Marks, 1987) and convey information about referents through an analog mapping of properties of those referents onto acoustic properties of speech. For example, speakers spontaneously modulate their fundamental frequency (F0, the acoustic correlate of pitch) when describing upward or downward motion, analogically mapping pitch height and vertical direction; speakers modulated their speech rate when describing fast or slow motion, analogically mapping rate of articulation to rate of motion (Shintel, Nusbaum, & Okrent, 2006). Analogically conveyed motion information influenced listeners' judgments about the described objects, even when information was conveyed exclusively in acoustic properties and not in the propositional content. Furthermore, such analog acoustic variation affected comprehension even when it was task irrelevant (Shintel & Nusbaum, 2007); listeners were faster to recognize that a picture represents an object mentioned in a preceding sentence when motion information implied in speech rate (fast vs. slow) matched the motion information implied in the picture (object in motion vs. at rest). These results suggest that analog acoustic variation can affect comprehension and that this effect reflects the natural part of comprehension.

However, it is not clear *how* analog variation in acoustic properties of speech affects comprehension. We can consider two alternative explanations. The first is that variation in acoustic properties affects listeners' responses or representations by some kind of stimulus-driven, automatic, context-independent priming process. For example, in our previous research, speech rate may have automatically primed speed-related concepts (e.g., fast speech activated the concept of speed) and thus influenced listeners' judgments about, or representation of, objects. In other words, the effect of analog acoustic variation may be determined by intrinsic acoustic properties of the stimulus itself; perception of fast-spoken utterances may have an automatic and invariant effect due to an acoustic property the utterance exemplifies.[1] Alternatively, the interpretation of speech rate may be context dependent in much the same way as other acoustic properties that convey linguistic and affective aspects of speech. A particular speech rate may convey different information in different contexts just as, for example, the same F0 difference can be interpreted as signalling a different talker or as reflecting random variability in different contexts (Magnuson & Nusbaum, 2007), the same F0 contour can be interpreted as conveying different affect in the context of different sentence types (Scherer, Ladd, & Silverman, 1984), and the same sentence-final rising intonation may indicate a question or surprise depending on listeners' expectations (Luks, Nusbaum, & Levy, 1998). On this account, the effect of a fast-spoken utterance is not a direct consequence of perceiving the fast speech rate, *per se,* but depends on the way speech rate is represented and interpreted in the course of communication.

One way to explore the context dependence of analog acoustic variation, specifically of speech rate, is by examining cases in which variation in speech rate may have an effect on listeners' actual behavior. Previous research supports the idea of cross-modal stimulus–response correspondences that can facilitate action (Kornblum, Habroucq, & Osman, 1990; for reviews, see Hommel & Prinz, 1997; Proctor & Reeve, 1990). Several findings have shown that stimuli can prime responses that share corresponding features. For example, participants responded

faster when response duration corresponded to the task-irrelevant stimulus duration (Kunde & Stöcker, 2002), or when response force corresponded to stimulus intensity (Mattes, Leuthold, & Ulrich, 2002; Romaiguère, Hasbroucq, Possamaï, & Seal, 1993). Our previous research suggests that, drawing on audio–visual cross-modal correspondences, acoustic properties of speech can provide referential information and affect comprehension. Similarly, drawing on stimulus–response correspondences, acoustic properties of speech may affect listeners' response; a fast or a slow speech stimulus may prime speed-corresponding responses.

In addition, considerable evidence shows that people often unconsciously match their behavior to the behavior of their interaction partner (Chartrand & Bargh, 1999; Dijksterhuis & Bargh, 2001). The tendency to match the behavior of one's interaction partner was observed across a wide range of speech-related behaviors such as accents (Giles, Coupland, & Coupland, 1991; Giles & Powesland, 1975), tone of voice (Neumann & Strack, 2000), speech rate (Webb, 1969, 1972), structural–syntactic choices (Branigan, Pickering, & Cleland, 2000), and referring expressions (Brennan & Clark, 1996; Garrod & Anderson, 1987). Listeners may adjust their response speed to the speaker's speech rate, even if they respond with an action rather than a verbal response.

Explanations of both unintentional mimicry and compatibility effects appealed to shared representations, or common coding, for perception and action as underlying the effect (Ferguson & Bargh, 2003; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Pickering & Garrod, 2004). Taken together, these findings raise the possibility that fast speech may prime a faster response. Such potential effect of perception on behavior may be driven exclusively by context-independent acoustic features of the stimulus. Alternatively, the effect of speech rate may depend on the context in which the utterance is presented. A fast-spoken utterance may be treated differently when presented in a context that supports an interpretation of fast speech rate as implying urgency compared to when the same utterance is presented in a context that does not support such interpretation. If the effect of speech rate on listener's response speed is modulated by context, listeners may respond faster when fast speech rate is treated as conveying relevant information, compared to when speech rate is not treated as conveying such information.

2. Experiment 1

To investigate these alternative hypotheses, we presented participants with short written scenarios. Some scenarios described a protagonist in a situation where he or she needs to quickly perform a specific action (typically involving pressing a button of some kind)—for example, submitting an online application before the deadline. Other scenarios did not have this implication—for example, there would be no mention of a strict deadline (relevant and irrelevant scenarios, respectively). See Table 1 for an example context scenario. Each scenario was followed by a recorded instruction to the participant to press different keyboard keys, spoken at a fast or a slow speech rate. Speech rate was never mentioned in the scenarios.

If speech rate can affect a listener's response speed, and if this effect reflects direct priming dependent only on stimulus properties, listeners should be faster to respond to fast-spoken

Table 1
Example context scenario in both conditions

Mark knew it was time to look for a new job and move on. Then one day he got a call from
his friend Dave. Dave was very excited about a great job ad he found and about how much he
hoped to get the job. The next day, Mark applied for the same job, and a few days later he was
scheduled for an interview. Mark felt bad about applying for the same job, and even worse for not
telling Dave. But the job was too good to give up and he thought that he may not get the job
anyway so Dave would never know. Mark's interview went very well. He was surprised by how
good he felt at the end of the interview.
Relevant ending:
*As he walked into the elevator, he saw a familiar person coming towards the elevator. He had to
go down before Dave could see him.*
Irrelevant ending (used only in Experiment 2):
*As he walked into the elevator, he decided to tell his family about the new job. He did not want to
tell anyone about the job before the interview.*

---

instructions, independently of the context. If, on the other hand, response speed depends not
only on speech rate in itself, but crucially on the information it carries in the context (e.g., a
fast-spoken instruction is more likely to be treated as implying urgency when preceded by a
scenario that highlights speed of action), we would expect a greater difference in listeners'
response speed to fast- and slow-spoken instructions following relevant scenarios compared
to irrelevant scenarios.

2.1. Method

2.1.1. Participants

Forty-nine University of Chicago students participated in the study. All had native fluency in English and no reported history of speech or hearing disorders. Participants were paid for their participation. Five participants were excluded from further analysis for not providing answers, or providing incorrect answers, to over 20% of the comprehension questions.

2.1.2. Materials

Twenty written scenarios served as experimental materials; each scenario appeared in a relevant version and an irrelevant version (never presented to the same participant). Relevant scenarios described the protagonist as being in a situation in which he needed to quickly perform a specific action. Irrelevant scenarios were constructed by omitting the last 1 or 2 sentences from relevant scenarios, so that they did not imply urgency (a list of scenarios used in the experiments is available at http://www.cogsci.rpi.edu/CSJarchive/Supplemental/index.html). There were 20 additional filler scenarios. Eight experimental and 8 filler scenarios were followed by true–false questions; each scenario was paired with a written sentence that participants had to judge as true–false based on the scenario.

Recorded sentences were produced by the same male speaker. Test instructions consisted of five imperative sentences, each constructed of the frame, "Press the — key," with a different keyboard key specified in each sentence. Each sentence was recorded spoken at a fast speech

rate (mean duration = 987 msec; 268 syllables per minute [SPM]) and at a slow speech rate (mean duration = 1,569 msec; 168 SPM), resulting in 10 test instructions. Five additional filler instructions were recorded at the speaker's natural speech rate (mean duration = 1,370 msec; 208 SPM).

2.1.3. Design and procedure

Relevance (relevant vs. irrelevant) and Instruction Speed (fast vs. slow) were manipulated within-subjects. Context scenarios were divided into two 20-sentence lists such that each list contained 10 relevant scenarios and the matching irrelevant scenarios. Each participant read 10 scenarios in each Relevance condition, taken from different lists (i.e., participants who received list A sentences for the relevant condition received list B sentences for the irrelevant condition), counterbalanced across participants. Of the 10 test sentences in each Relevance condition, 5 were paired with a fast instruction and 5 with a slow instruction, counterbalanced across participants.

Participants read each scenario and were instructed to press the spacebar when they were done. Immediately after pressing the spacebar, participants heard a recorded instruction and responded by pressing a key. The spoken instruction was not part of the scenario, although the action that listeners were instructed to perform (pressing a key) did bear a resemblance to the action the protagonist was about to perform in the action-relevant scenarios (e.g., ringing a doorbell).

Context scenarios were blocked by Relevance condition, each block containing 10 experimental and 10 filler scenarios, presented in a random order. Block order was counterbalanced across participants. To ensure participants indeed read the stories, 8 trials in each block (4 experimental and 4 filler) were followed by a true/false question. On other trials, participants saw a sentence instructing them to press the spacebar to continue to the next trial. Response times were measured from the phonetic point of disambiguation of the critical key name (cf. Grosjean, 1980; Marlsen Wilson, 1984), defined as the point when the word could have referred to a unique key relative to other instructions (e.g., the second phoneme in "e<u>n</u>d" and "e<u>s</u>cape").

Because relevant scenarios ended with a sentence describing or implying a reaching action, they may have primed a corresponding motor response (Glenberg & Kaschak, 2002). Because the time intervals between scenario offset and critical word onset differed for fast and slow instructions (short and long intervals, respectively), this may have resulted in different degrees of motor activation decay in the two Speed conditions. Such a potential difference in motor priming may be reflected in a response time difference between fast and slow instructions in the relevant condition, but not in the irrelevant condition because irrelevant scenarios did not prime a motor response. To evaluate if the effect is due to the time interval, *per se* (rather than to speech rate), 8 participants participated in a control condition with written instructions. In each Speed condition, time intervals between scenario offset and written instruction onset matched the intervals between scenario offset and spoken instruction disambiguation point. In all other respects, the procedure was the same. If the effect is due to the different time intervals, a similar pattern of results should emerge in both the spoken and the written versions.
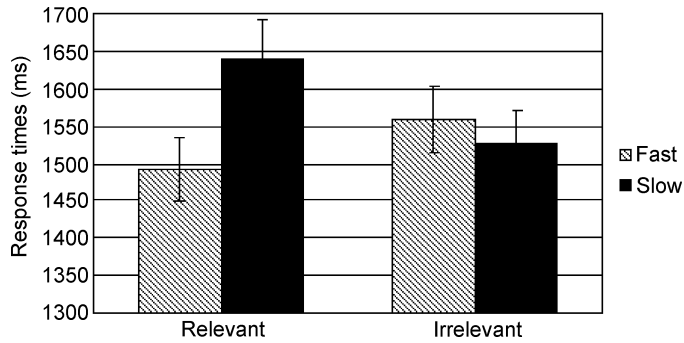
Fig. 1. Response times by condition in Experiment 1.

2.2. Results and discussion

Trials followed by an incorrect key press or an incorrect answer on the true/false question were excluded from the analysis (<4% of test trials). Key errors were not further analyzed because of low error rate (< 1%). Response times greater than 3,500 msec, or greater than 2.5 *SD*s above the participant's subsequent mean in the Speed condition, were excluded from the analysis (<3% of test trials).

A repeated-measures analysis of variance (ANOVA) conducted on participants' response times in the spoken version showed a significant Relevance × Instruction Speed interaction, $F_1(1, 35) = 7.9, MSE = 36,826, p < .01$; marginal in the item analysis, $F_2(1, 19) = 4.27, MSE = 20,116, p < .053$. As shown in Fig. 1, response times were shorter for fast instructions than for slow instructions in the relevant condition (1,492 and 1,640 msec, respectively); however, the reverse pattern emerged in the irrelevant condition (1,560 and 1,527 msec, respectively). A simple effects analysis revealed a significant difference between fast instructions and slow instructions in the relevant condition, $t_1(35) = 3.2, p < .005; t_2(19) = 2.67, p < .02$. The difference was not significant in the irrelevant condition: all ts < 1 (in fact, numerically, the difference in the irrelevant condition was in the direction opposite to the relative speech rate). The main effect of speed was marginal, $F_1(1, 35) = 3, MSE = 39,734, p < .1; F_2(1, 19) = 2.81, p > 1, ns$. Relevance was not significant (all Fs < 1), suggesting the effect is not due to discourse context alone.

However, response times in the written control condition revealed no significant effects (all Fs < 1, *ns*). There was no difference in response times for short and long intervals in the relevant condition (1,709 and 1,707 msec, respectively; $p > .9, ns$). This pattern of results argues that the difference between fast and slow instructions following relevant scenarios depends on speech rate, rather than merely reflecting the time course motor activation decay.

The written control condition also provides some baseline for evaluating the effect of context scenarios independently of the effect of speech rate. Although the difference was not significant (perhaps due to the small number of participants), response times were longer following relevant stories than following irrelevant stories (1,708 and 1,615 msec, respectively). One possibility is that participants devoted more processing resources to relevant stories, perhaps engaging in more elaborative inferences. Thus, participants may have had fewer resources

available for responding to the instructions, resulting in longer response times. Critically, this potential difference did not interact with the time interval between scenario offset and instruction onset. Response times in the spoken version should be interpreted against the background of this difference. Although response times in the relevant context/fast instructions cell in the spoken version did not differ significantly from response times in the irrelevant context/slow instructions cell, the two Relevance conditions have different baselines and thus cannot be strictly compared.

The significant Relevance × Speed interaction argues that the effect of speech rate on listeners' response is not determined exclusively by the intrinsic acoustic property of speech rate. Instead, the results suggest that the effect of speech rate may depend on the way listeners represent and interpret speech rate information as derived from understanding the antecedent discourse. Listeners responded faster to fast instructions when these followed contexts that emphasized urgency and speed of action. However, in the absence of such contextual information, following discourse with no temporal content or implications, variation in speech rate may have been treated as reflecting normal random articulatory variability and thus had no reliable effect on subsequent response speed.

Because relevant contexts suggested urgency, fast-spoken instructions could be viewed as matching the context, whereas slow-spoken instructions mismatch the context. It is possible that faster responses for fast instructions resulted from a match between the stimulus and the context rather than from fast speech priming specific response tendencies. Although this is possible, such context-stimulus matches cannot constrain the possible responses, as there was no relation between the context and the specific key indicated. Thus, the context could not have facilitated specific responses in "matching" trials. Such facilitation could occur if, in every case, listeners were trying to relate the speech rate to the (supposedly unrelated) context before responding. In post-experimental questioning, only 3 participants mentioned variations both in speech rate and scenario urgency, suggesting participants were not consciously trying to relate speech rate to the scenarios. It is interesting to note that these participants showed the reverse pattern; they responded slower to fast-spoken compared to slow-spoken instructions in both the relevant (1,673 vs. 1,521 msec, respectively) and the irrelevant (1,502 vs. 1,425 msec, respectively) conditions. Although listeners may try to relate speech rate to the context subconsciously, this explanation, as the previous one, is based on the idea that the interpretation of speech rate is context dependent and that its effect is not determined exclusively by an invariant mapping between an acoustic property and a response.

However, it is possible that participants implicitly recognized a speed-related theme in the scenarios and subconsciously adopted a general strategy of attending to speech rate. Given that instructions were blocked by Relevance condition, participants may have adopted this strategy in the relevant block, but not in the irrelevant block. If this is the case, the response speed difference in the relevant condition may reflect a task-specific response pattern, learned over the course of the task. On the other hand, if the effect indeed reflects a non-strategic interaction between speech rate and discourse context that constrains the response to the current trial, we would expect the same pattern of results when both types of context scenarios are mixed.

A further issue concerns the potential acoustic differences between fast and slow instructions. Because utterances were naturally produced, they may have differed on acoustic and phonetic dimensions other than speech rate. Although the same utterances were used in both

Relevance conditions (and, hence, the effect cannot be due to acoustic information alone), such a potential difference makes it difficult to isolate the specific acoustic property underlying the effect. These issues are addressed in the next experiment.

3. Experiment 2

To ensure that the difference between the Relevance conditions did not reflect response strategies that resulted from the blocked presentation of the stimuli, in Experiment 2, trials were presented in a random order instead of blocked by Relevance condition.

Furthermore, to examine whether speech rate difference was indeed the critical acoustic difference underlying the effect found in Experiment 1, rate was manipulated synthetically instead of using naturally produced fast and slow spoken instructions. In this way, we can be sure that the two Instruction Speed conditions did not differ on other acoustic/phonetic parameters.

3.1. Method

3.1.1. Participants

Fifty-eight University of Chicago students participated in the study. All had native fluency in English and no reported history of speech or hearing disorders. Participants were paid for their participation. Two participants were excluded from further analysis due to greater than 20% error rate on test trials or extremely high response times (over 3 *SD*s above the mean for all participants).

3.1.2. Materials

Twenty-four short written scenarios served as the experimental materials; each appearing in the two Relevance versions as in Experiment 1. An additional sentence, unrelated to speed or to a to-be-performed action was added to irrelevant scenarios to make scenario length comparable across conditions. There were an additional 36 filler scenarios. To encourage participants to read the scenarios, all scenarios were followed by true/false comprehension questions.

Six recorded sentences were produced by a male speaker (mean duration = 1,198 msec; 209 SPM). Speech rate was then manipulated synthetically using the PSOLA algorithm in the Praat software (Boersma & Weenink, 2006), resulting in a fast version (mean duration = 896 msec; 279 SPM) and a slow version (1,454 mec; 172 SPM) of each instruction. Eighteen filler instructions (9 sentences, each recorded 2 times) were recorded at the speaker's natural speech rate (1,370 msec; 216 SPM).

3.1.3. Design and procedure

To familiarize participants with the keyboard keys used for responding, participants first completed a practice task that required them to respond by pressing different keys. Because we did not want to mention response speed, so as not to encourage participants to infer anything about the required response speed based on the practice task, we used a memory task in which

names of different keyboard keys appeared sequentially on the screen, followed by a square indicating the number "1" or "2." Participants responded by pressing the key that preceded the square by 1 or 2 positions. Items included the keys used in the critical instructions, as well as keys included in the filler instructions or not included in the experimental phase.

The test procedure was the same as in Experiment 1, except that stimuli were presented in a random order and all trials were followed by questions. Response times were measured from the onset of the critical key name (key names did not share an initial phoneme).

3.2. Results and discussion

Trials followed by an incorrect key or an incorrect answer on the true/false question were excluded from the analysis (4.2% of test trials). Key errors were not further analyzed because of low error rate (<1%). Response times greater than 3,500 msec, or greater than 2.5 SDs above the participant's subsequent mean, were excluded from the analysis (<3% of test trials).

ANOVA of the response times revealed a significant effect of Instruction Speed—$F_1(1, 55) = 16.18$, $MSE = 17,111$, $p < .001$; $F_2(1, 23) = 20.97$, $MSE = 6,484$, $p < .001$—such that response times were faster for fast, compared to slow, instructions. However, this effect was qualified by a significant Relevance by Instruction Speed interaction: $F_1(1,55) = 6.76$, $MSE = 13,186$, $p < .02$ (not significant by items); $F_2(1, 23) = 1.65$, $MSE = 11,833$, $p < .22$. As shown in Fig. 2 the difference between response times for fast and slow instructions was greater in the relevant condition (1,310 and 1,420 msec, respectively), compared to the irrelevant condition (1,325 vs. 1,355 msec, respectively). Although the interaction was not significant by items, this was probably due to low statistical power given the small number of items. Although a larger number of items would be desirable in principle, given the specific theme common to all the test scenarios, increasing the number of items was likely to increase participants' awareness of the experimental manipulation. Therefore, we limited the number of items in order to ensure participants are not aware of the manipulation. More important, the majority of the items did show the expected pattern (17 out of 24, significant in a one-tailed sign test, $p < .04$), suggesting the effect was not due to a few distinctive items. Furthermore, a simple effects analysis revealed that, in both the participants and the items analyses, the Speed difference was significant only in the relevant condition, $t_1(55) = 4.31$, $p < .001$ and
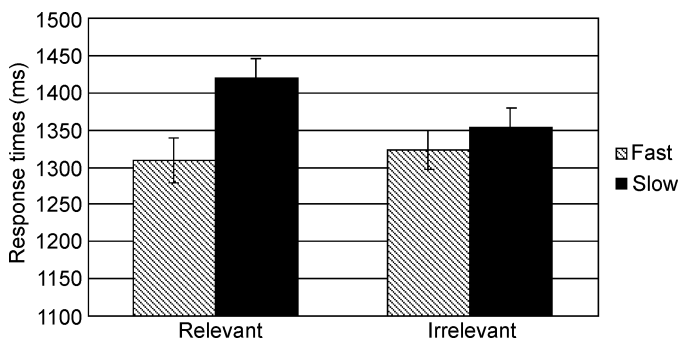

Fig. 2. Response times by condition in Experiment 2.

$t_2(23) = 3.75$, $p = .001$; but not in the irrelevant condition, $t_1(55) = 1.47$ $p > .1$ and $t_2(23)$ $= 1.7$, $p > .1$. Finally, estimation of effect size by items revealed a similar pattern with a bigger effect in the relevant condition ($d = .54$, adjusted for repeated-measures design; see Dunlap, Cortina, Vaslow, & Burke, 1996), compared to the irrelevant condition ($d = .21$). These results show that the pattern observed in Experiment 1 does not reflect a task-specific learning effect or a strategy of attending to speech rate that is applied in a global manner to the entire block. Instead, it appears that the interaction of speech rate and discourse context constrained listeners' responses on a trial by trial basis, and that listeners were sensitive to local changes in the context. The effect of Relevance was marginal only by participants, $F_1(1, 55) = 3.22$, $MSE = 10,730$, $p < .1$; $F_2 < 1$, *ns*.

These results replicate the results of Experiment 1, and further support the idea that speech rate has an effect on listeners' behavior; a faster speech rate led to faster responses. However, this effect is constrained by the discourse context rather than exclusively determined by acoustic properties of the speech. The effect of context suggests that the physical or perceptual properties of the stimulus do not fully control listeners' response tendencies. Rather, the effect, although not contingent on listeners' awareness or on a voluntary decision stage, is modulated by the way the stimulus is represented and interpreted. Such an interpretation is consistent with findings suggesting that even seemingly straightforward spatial stimulus–response compatibility effects, such as the Simon effect, can be modulated by task instructions, participants' intentions, stimulus context, and the way people attend to and code both stimulus and response (Guiard, 1983; Hommel, 1993; Hommel & Lippa, 1995; see also Hommel et al., 2001).

The present results extend our previous studies of analog variation in acoustic properties of speech, demonstrating that speech rate can affect listeners' behavior, as well as their representation of described objects. Although we have focused here on speech rate, research on stimulus–response compatibility (e.g., Mattes et al., 2002; Romaiguère et al., 1993) suggests that other acoustic properties of speech may affect listeners (e.g., stimulus amplitude may affect response force). However, our findings also suggest that the effect of speech rate on listeners' response is not determined just by acoustic properties of the stimulus in itself, but depends on the context in which those properties are presented. More broadly, these findings suggest that acoustic properties of speech that have generally been thought to be irrelevant from a linguistic point of view can be relevant for spoken communication.

Note

1. Speech rate may be evaluated relative to the speaker's speech rate in other utterances, and in this sense is not strictly an intrinsic acoustic property of the utterance in itself.

Acknowledgments

References

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science, 15,* 578–582.

Bailey, K. G. D., & Ferreira, F. (2003). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory & Language, 49,* 183–200.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality & Social Psychology, 70,* 614–636.

Barr, D. J. (2003). Paralinguistic correlates of conceptual structure. *Psychonomic Bulletin & Review, 10,* 462–467.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory & Language, 30,* 644–663.

Birch, S., & Clifton, C., Jr. (1995) Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech, 33,* 365–391.

Bock, K., & Mazzella, J. R. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition, 11,* 64–76.

Boersma, P., & Weenink, D. (2006). *Praat: Doing phonetics by computer* (version 4.4.32) [Computer program]. Retrieved September 30, 2006, from http://www.praat.org

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic coordination in dialogue. *Cognition, 75,* B13–B25.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22,* 1482–1493.

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

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality & Social Psychology, 76,* 893–910.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84,* 73–111.

Dijksterhuis, A., & Bargh, J. A. (2001). The perception–behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 1–40). San Diego, CA: Academic.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1,* 170–177.

Ferguson, M. J., & Bargh, J. A. (2003). How social perception can automatically influence behavior. *Trends in Cognitive Sciences, 8,* 33–39.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27,* 181–218.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In G. Howard & J. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics. Studies in emotion and social interaction* (pp. 1–68). New York: Cambridge University Press.

Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation*. Oxford, England: Academic.

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

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28,* 267–283.

Guiard, Y. (1983). The lateral coding of rotations: A study of the Simon effect with wheel-rotation responses. *Journal of Motor Behavior, 15,* 331–342.

Hommel, B. (1993). Inverting the Simon effect by intention. *Psychological Research, 55,* 270–279.

Hommel, B., & Lippa, Y. (1995). S–R compatibility effects due to context-dependent spatial stimulus coding. *Psychological Bulleting and Review, 2,* 370–374.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24,* 849–878.

Hommel, B., & Prinz, W. (Eds.). (1997). *Theoretical issues in stimulus–response compatibility.* Amsterdam: North-Holland.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility—A model and taxonomy. *Psychological Review, 97,* 253–270.

Kunde, W., & Stöcker, C. (2002). A Simon effect for stimulus–response duration. *The Quarterly Journal of Experimental Psychology, 55A,* 581–592.

Luks, T. L., Nusbaum, H. C., & Levy, J. (1998). Hemispheric involvement in the perception of syntactic prosody is dynamically dependent on task demands. *Brain and Language, 65,* 313–332.

Magnuson, J. S., & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception & Performance, 33,* 391–409.

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception & Performance, 13*(3), 384–394.

Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance: Control of language processes* (pp. 125–150). Hillsdale, NJ: Erlbaum.

Mattes, S., Leuthold, H., & Ulrich, R. (2002). Stimulus–response compatibility in intensity–force relations. *The Quarterly Journal of Experimental Psychology, 55A,* 1175–1191.

Neumann, R., & Strack, F. (2000). "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality & Social Psychology, 79,* 211–223.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral & Brain Sciences, 27,* 169–226.

Proctor, R. W., & Reeve, T. G. (Eds.). (1990). *Stimulus–response compatibility: An integrated approach.* Amsterdam: Elsevier.

Romaiguère, P., Hasbroucq, T., Possamaï, C. A., & Seal, J. (1993). Intensity to force translation: A new effect of stimulus–response compatibility revealed by analysis of response time and electromyographic activity of a prime mover. *Cognitive Brain Research, 1,* 197–201.

Scherer, K. R., Ladd, D. R., & Silverman, K. A. (1984). Vocal cues to speaker affect: Testing two models. *Journal of the Acoustical Society of America, 76,* 1346–1356.

Shintel, H., & Nusbaum, H. C. (2007). The sound of motion in spoken language: Visual information conveyed by acoustic properties of speech. *Cognition, 105,* 681–690.

Shintel, H., & Nusbaum, H. C. (2007). Speed accommodation in context: Context modulation of the effect of speech rate on response speed. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 641–645). Austin, TX: Cognitive Science Society.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech. *Journal of Memory & Language, 55,* 167–177.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory & Language, 48,* 103–130.

Terken, J., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes, 2,* 145–163.

Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes, 19,* 713–755.

Webb, J. T. (1969). Subject speech rates as a function of interviewer behaviour. *Language and Speech, 12,* 54–67.

Webb, J. T. (1972). Interview synchrony: An investigation of two speech rate measures in an automated standardized interview. In B. Pope & A. W. Seigman (Eds.), *Studies in dyadic communication* (pp. 115–133). New York: Pergamon.