

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Speed Accommodation in Context: Context Modulation of the Effect of Speech Rate on Response Speed

Permalink

<https://escholarship.org/uc/item/4vc028rw>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN

1069-7977

Authors

Shintel, Hadas
Nusbaum, Howard C.

Publication Date

2007

Peer reviewed

Speed Accommodation in Context: Context Modulation of the Effect of Speech Rate on Response Speed

Hadas Shintel (hadas@uchicago.edu)

Department of Psychology and Center for Cognitive and Social Neuroscience,
The University of Chicago
Chicago IL 60637 USA

Howard C. Nusbaum (h-nusbaum@uchicago.edu)

Department of Psychology and Center for Cognitive and Social Neuroscience,
The University of Chicago
Chicago IL 60637 USA

Abstract

Previous research has shown that analog mapping of information onto acoustic properties of speech can affect listeners' interpretation of described objects and events. In this paper, we examine whether the effect of analog acoustic variation on listeners is governed by intrinsic acoustic properties of the stimulus, through direct priming of listeners' concepts or behavior. For example, fast speech may always prime the concept of speed or a faster response. Alternatively, the effect of analog acoustic variation may be mediated by a context-dependent interpretation of the acoustic properties. For example, the effect of speech rate may depend on how listeners construe the meaning of a fast or a slow speech rate in the immediate linguistic context. Our results suggest that speech rate can have an effect on listeners' response speed, but only in contexts that supported the interpretation of speech rate as conveying response speed relevant information.

Keywords: Prosody; Perception- behavior link; Language comprehension.

Introduction

Numerous studies have shown that prosodic patterns (see Cutler, Dahan, & Van Donselaar, 1997, for a review) can be meaningful to listeners and systematically affect comprehension. Recently, we have shown that, in addition to conveying information about the speaker's internal state (such as emotion or certainty) or the message (syntactic structure or discourse status), prosody can convey referential information through an analog mapping of information about object properties onto acoustic properties of speech (Shintel, Nusbaum, & Okrent, 2006). For example, speakers spontaneously modulated their fundamental frequency (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. Such analog acoustic information affected comprehension even when it was irrelevant for listeners' task and was not correlated with task-relevant information (Shintel & Nusbaum, in press); listeners were faster to

recognize that a picture represents an object that had been 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). Acoustically-conveyed motion information affected listeners even though it was not required for the recognition task, and even though its use did not confer any performance benefit.

These findings suggest that, capitalizing on audio-visual cross-modal correspondences, acoustic properties of speech can provide non-auditory referential information. While these experiments focused on a referential function of cross-modal analog mappings of information onto acoustic properties of speech, it is possible that such analog mappings can also have a non-referential function.

Previous research has supported the idea of cross-modal stimulus-response correspondences and has shown that auditory or visual stimuli can prime responses that share a compatible feature, for example corresponding duration or intensity (Kunde & Stöcker, 2002; Mattes, Leuthold, & Ulrich, 2002; Romaiguère, Hasbroucq, Possamaï, & Seal, 1993). These findings raise the possibility that, drawing on stimulus-response correspondences, acoustic properties of speech may also affect listeners' response, not just their representations of external referents. For example a fast or a slow speech stimulus may prime speed-corresponding responses.

In addition to stimulus-response compatibility effects, considerable evidence suggests that people often unconsciously mimic each other or 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 conversational partner was observed across a wide range of speech-related behaviors such as accents (Giles & Powlsland, 1975; Giles, Coupland, & Coupland, 1991), tone of voice (Neumann & Strack, 2000), speech rate (Webb 1969; Webb 1972), as well as across other linguistic levels (see Pickering & Garrod, 2004). Although the effect of the perception of a stimulus on

subsequent behavior may be modulated by the degree of similarity between the perceived stimulus and the relevant behavior, speed accommodation may occur even when listeners respond to a speech stimulus with a non-verbal behavior.

A potential effect on listeners' behavior also raises the issue of the process underlying the effect of acoustically conveyed analog information. Explanations of compatibility phenomena, as well as unintentional mimicry, have called upon the idea of shared representations, or common coding, for perception and action (see Chartrand & Bargh, 1999; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Pickering & Garrod, 2004). However, perception may affect behavior through different processes. One possibility is that stimulus features prime response features through a direct, automatic, context-independent priming process. For example, in our previous studies a fast speech rate may directly prime the concept of a fast speed, and in this way prime a representation of the object as being in motion. Thus, the effect of analog acoustic variation may be determined by intrinsic acoustic properties of the stimulus itself. An alternative possibility is that the effect of a fast-spoken utterance is not a direct consequence of perceiving fast speech rate per se, but is mediated by a context-dependent interpretation of the modulation of speech rate; a particular speech rate may convey different information in different contexts and have different effects on listeners' representations and responses.

The current experiment

To investigate these hypotheses, we presented participants with written short scenarios. Some scenarios described the protagonist in a situation that required him/ her to quickly perform a specific action, for example submitting an online application right before the deadline, thereby making speed-related information relevant. Other scenarios did not have this implication, for example there would be no mention of a strict deadline approaching (relevant and irrelevant scenarios respectively). Each scenario was followed by a recorded instruction to the participant to press different keyboard keys, spoken at a fast or at a slow speech rate.

If speech rate affects listeners' response speed, and if this effect reflects direct priming due to the perception of intrinsic properties of the stimulus, listeners should be faster to respond to fast-spoken instructions regardless of the context; there should be no interaction between the speech rate with which the instruction is given and the type of context scenario that preceded it. On the other hand, if listeners' response speed depends not only on the speech rate in itself, but on the information it carries given the immediate context (e.g. a fast spoken instruction is more likely be treated as implying urgency when it is preceded by a scenario in which speed of action is made highly relevant), we would expect listeners to respond faster to fast-spoken instructions compared to slow-spoken instructions following

the relevant scenarios, but not following the irrelevant scenarios.

Context scenario in both Relevance conditions (parts in bold appeared only in the relevant condition)
James was very proud of his daughter. As a child she was exceptionally bright, but even he and her mother were surprised when she finished her PhD at such a young age and went on to become a neurology professor at a distinguished university. James was particularly excited when the results of her last study, investigating the neurotransmitters involved in the brain's recovery following a stroke, were reported in the science section of the New York Times. Not only that, but she was about to be interviewed on the 7 o'clock evening news on TV. She told him it would only be a short interview, no longer than a couple of minutes. James made sure he was home to watch the interview. Suddenly he noticed that it was already 7 and he reached for the remote control to turn on the TV.

Figure 1: An example of one experimental scenario

Method

Participants

45 University of Chicago students contributed usable data (37 participated in the experiment and 8 in a control conditions, see Design and Procedure). All participants had native fluency in English and no reported history of speech or hearing disorders.

Materials

Twenty short scenarios served as the experimental materials. Each scenario appeared in two versions: relevant scenarios described the protagonist as being in a situation where she/ he needed to quickly perform a specific action, (typically involving pressing some kind of button, for example ringing a doorbell or pressing an elevator button), irrelevant scenarios were constructed by omitting the last 1-2 sentences from each relevant scenario, so that they did not imply a need to quickly perform an action. Twenty additional scenarios served as filler scenarios.

Recorded sentences were produced by a male speaker. Test instructions consisted of five imperative sentences, each with a different keyboard key name (e.g. *End* or *Esc*) embedded in the phrase *Press the ___ key*. Each sentence was recorded by a male speaker spoken at a fast speech rate (mean duration 987 ms) and at a slow speech rate (mean duration 1569 ms), resulting in ten test instructions. Reaction time was measured from the disambiguation point of the key name relative to other key names (e.g. the second phoneme in *End* and *Esc*). To make the experimental manipulation less obvious, five additional filler instructions

were recorded at the speaker's natural speech rate (mean duration 1370 ms).

Design and procedure

The design consisted of two levels of Relevance (relevant vs. irrelevant) crossed with two levels of Instruction Speed (fast vs. slow) manipulated within-subjects. Each participant read ten scenarios in each Relevance condition (scenario versions were counterbalanced across participants). Out of the experimental scenarios in each Relevance condition, half were paired with a fast instruction and half were paired with a slow instruction, counterbalanced across participants.

Scenarios were presented on a computer screen. Participants were instructed to read each scenario at their own pace and to press the spacebar when they were done reading. Immediately after pressing the spacebar, a spoken instruction was presented through headphones and participants had to respond by pressing the appropriate keyboard key. Eight experimental trials and 8 filler trials were followed by a true/ false question; after responding to the instruction, participants saw a sentence and had to indicate whether it was true or false in the scenario. Context scenarios were blocked by Relevance condition, each block containing ten experimental scenarios and ten filler scenarios. Block order was counterbalanced across participants.

Because relevant context scenarios ended with a sentence describing or implying a reaching/ pressing action, reading the sentence may have primed a corresponding motor response (see Glenberg & Kaschak, 2002). Since the time intervals between the scenario offset and the critical word onset differed for fast and slow instructions (short or long intervals, respectively), this may have resulted in different degrees of decay of the motor activation for the two conditions, and thus in different effects on listeners' response speed¹. To ensure that the difference in response speed is due to speech rate, rather than to the time interval per se, 8 participants participated in a 'no-speech' control condition with written instructions. In each speed condition, time intervals between scenario offset and written instruction onset matched the time intervals between scenario offset and spoken instruction disambiguation point. In all other respects, the procedure was the same as in the spoken version. If the difference in response speed is due to the different time interval, a similar pattern of results should emerge in both the spoken and the written versions.

Results and Discussion

Trials followed by an incorrect key press or an incorrect answer on the true/ false question were excluded from the analysis (3.7% of all trials; <1% in the control condition). In addition, response times greater than 3500 ms, or greater than 2.5 standard deviations above the participant's mean in

that Instruction Speed condition, were excluded from the analysis (2.4% of all trials; 4.4% in the control condition).

A repeated-measures ANOVA conducted on participants' response times showed that there was no significant main effect of Relevance ($F_1(1,36) < 1$, $F_2(1,19) = 1.4$, $p_s > .25$, NS). However, there was a marginally significant effect of Instruction Speed ($F_1(1,36) = 2.7$, $MSE = 43141$ $p < .11$, non significant in the item analysis, $F_2(1,19) < 1$). This marginal difference was qualified by a significant Relevance by Instruction Speed interaction ($F_1(1,36) = 9.47$, $MSE = 35286$ $p_1 < .005$; $F_2(1,19) = 5$, $MSE = 30501$, $p_2 < .05$). Response times were shorter for fast instructions compared to slow instructions in the relevant condition (1482 ms. and 1633 ms. respectively). However, the reverse pattern emerged in the irrelevant condition (1551 ms. and 1513. respectively). A simple effects analysis showed that the difference in response times between fast instructions and slow instructions was significant in the relevant condition, $t(36) = 3.5$, $p < .005$, but not in the irrelevant condition, $t < 1$, NS.

A repeated-measures ANOVA conducted on the response times of participants in the control version with the written instructions showed no significant effects (all $F < 1$, NS). Within the relevant condition, response times for fast and slow instructions (i.e. short and long intervals, respectively) did not differ (1709 and 1707 ms, respectively, $p > .9$, NS). The results thus support the idea that the difference between fast and slow spoken instructions following relevant contexts scenarios in the spoken version of the experiment depends on speech rate, rather than merely reflecting the time course motor activation decay.

These results show that speech rate had an effect on the speed with which listeners performed the required response. However, the effect of speech rate was different depending on the type of the preceding context. The interaction of speech rate with the preceding context shows that the effect on listeners' response speed is not governed only by the speech rate of the stimulus, and thus argues against the idea that listeners' response speed is directly primed by the intrinsic acoustic property of the stimulus, independently of the context.

In post-experimental questioning participants were asked about the goal of the study, the goal of the speaker producing the instructions, and if they noticed anything special about the stories or the instructions. Only 4 participants mentioned both speech rate and implied urgency in the scenarios, suggesting that participants were not generally aware of the relation between speech rate and the context scenario. These 4 participants actually showed an opposite pattern; they responded slower to fast-spoken compared to slow-spoken instructions in both the relevant (1537 vs. 1438 ms, respectively) and the irrelevant (1441 vs. 1342) conditions.

¹ We thank an anonymous reviewer for this suggestion.

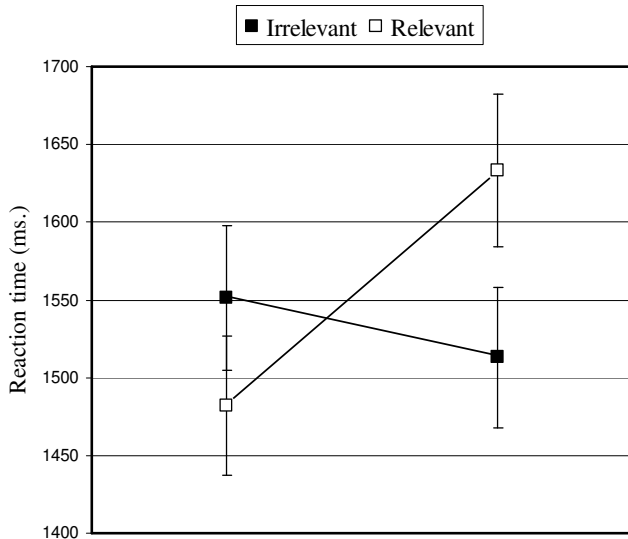


Figure 2: Reaction times by Relevance condition

The results support the idea that speech rate can have a reliable effect on listeners' response speed. The results further suggest that the effect of speech rate on listeners' responses may depend on the way listeners interpret speech rate information, based on the antecedent discourse. Listeners responded faster to fast instructions when these followed contexts that made speed of action and speed information highly relevant and supported an interpretation of fast speech as signaling a need for a faster action. However, speech rate had no effect on listeners' response speed in the absence of such contextual information. Context scenarios with speed-related content may have primed the dimension of speed and called attention to speed information as relevant information in the context. On the other hand, variation in speech rate following contexts that had no timing implications or speed-related content may have been treated as reflecting normal random articulatory variability and thus have no reliable effect on listeners' subsequent response speed. Such context-dependent interpretation of speech rate is consistent with findings regarding the context-dependent perception and interpretation of variations in acoustic properties of speech that convey linguistic or affective information. For example, depending on listeners' expectations about the speech presented to them, the same fundamental frequency (F0) difference was interpreted as signaling a different talker or as reflecting random variability (Magnuson & Nusbaum, in press), and the same sentence final F0 rise was interpreted as indicating a question or surprise (Luks, Nusbaum, & Levy, 1998). Similarly, the same F0 contour was interpreted as conveying different speaker affect in the context of different sentence types (Scherer, Ladd, & Silverman, 1984).

These results suggest that while the effect of speech rate on listeners' response is not contingent on listeners' conscious intention to respond at a specific speed, it may be modulated by the way listeners code and attend to speech rate in specific contexts. Such an interpretation would be consistent with findings on stimulus-response compatibility effects, such as the Simon effect (Simon, 1969), in which response time to non-spatial stimulus features is affected by the spatial compatibility between the task-irrelevant location of the stimulus and the location of the response. Previous findings suggest that even such spatial compatibility effects that involve a seemingly direct relation between stimulus and response can be modulated by task instructions, participants' intentions, stimulus context, and the way people code and interpret the stimulus and the response. For example, Hommel (1993) has shown that when a key-press response was associated with a light flashing on the opposite location (e.g. a right hand key switched the left light), the direction of the Simon spatial compatibility effect depended on whether participants coded the same response as 'key pressing' or as 'light flashing'. Similarly, Guiard (1983) has shown that when turning a steering wheel to the right required moving both hands to the left, the spatial compatibility effect was modulated by whether participants coded the same response as hand movement or as wheel rotation. These findings emphasize a relation between perception and behavior that does not require an intermediary voluntary translation process, but nonetheless is modulated by relevance, context, and goals, rather than consisting of an invariant mapping process between specific perceptual features of the stimulus and specific motor features of the response (for example see the theory of event coding, Hommel et al., 2001).

The present results suggest that acoustic properties of speech that may not appear relevant from a linguistic point of view, for example speech rate, may nonetheless have an effect on listeners and may affect listeners' behavior, as well as their interpretation of the described objects and events. This effect may not be limited to speech rate, but may extend to other acoustic properties of speech. For example, previous studies showed a stimulus-response compatibility effect between the stimulus and response duration (Romaiguère et al., 1993) and between stimulus intensity and response force (Mattes et al., 2002), suggesting that other acoustic properties, such as amplitude, may have an effect on listeners response.

However, our findings also suggest that the effect of speech rate on listeners' response is not governed just by the speech stimulus in itself, but depends on the stimulus context and on the way the acoustic properties of the stimulus are interpreted within that context.

Acknowledgments

We thank Tim Brawn for his help with stimulus preparation. The support of the Center for Cognitive and Social Neuroscience at The University of Chicago is gratefully acknowledged.

References

- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality & Social Psychology*, 76(6), 893-910.
- Cutler, A., Dahan, D., & Van Donselaar, W.A. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40 (2), 141-202.
- 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). San Diego: Academic Press.
- 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*. New York, NY: Cambridge University Press.
- Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation*. Oxford, England: Academic Press.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558-565
- Guiard, Y. (1983). The lateral coding of rotations: A study of the Simon effect with wheel-rotation responses. *Journal of motor behavior*, 15(4), 331-342.
- Hommel, B. (1993). Inverting the Simon effect by intention. *Psychological Research*, 55, 270-279.
- 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.
- Kunde, W. & Stöcker, C. (2002). A Simon effect for stimulus-response duration. *The Quarterly Journal of Experimental Psychology*, 55A(2), 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. (In Press). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception & Performance*.
- Mattes, S., Leuthold, H., & Ulrich, R. (2002). Stimulus-response compatibility in intensity-force relations. *The Quarterly Journal of Experimental Psychology*, 55A(4), 1175-1191.
- Neumann, R., & Strack, F. (2000). "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality & Social Psychology*, 79(2), 211-223.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral & Brain Sciences*, 27(2), 169-226.
- 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. (In press). The sound of motion in spoken language: Visual information conveyed by acoustic properties of speech. *Cognition*.
- Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech. *Journal of Memory and Language*, 55, 167-177.
- Simon, H. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81, 174-176
- 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*. New York: Pergamon.