Phonetic categorization and discrimination of voice onset time under divided attention

Faith Chiu, Lyndon L. Rakusen, and Sven L. Mattys

Phonetic categorization and discrimination of voice onset time under divided attention

Faith Chiu,[1,a)] **Lyndon L. Rakusen,**[2] **and Sven L. Mattys**[2,b)]

[1]*Department of Language and Linguistics, University of Essex, Colchester, United Kingdom*
[2]*Department of Psychology, University of York, York, United Kingdom*
f.chiu@essex.ac.uk, lyndon.rakusen@york.ac.uk, sven.mattys@york.ac.uk

Abstract: Event durations are perceived to be shorter under divided attention. "Time shrinkage" is thought to be due to rapid attentional switches between tasks, leading to a loss of input samples, and hence, an under-estimation of duration. However, few studies have considered whether this phenomenon applies to durations relevant to time-based phonetic categorization. In this study, participants categorized auditory stimuli varying in voice onset time (VOT) as /g/ or /k/. They did so under focused attention (auditory task alone) or while performing a low-level visual task at the same time (divided attention). Under divided attention, there was increased response imprecision but no bias toward hearing /g/, the shorter-VOT sound. It is concluded that sample loss under divided attention does not apply to the perception of phonetic contrasts within the VOT range.
© 2020 Acoustical Society of America

1. Introduction

Executing a task under divided attention often leads to a decrease in performance (e.g., Navon and Gopher, 1979). The same holds for speech perception, with phonetic discrimination being impaired under cognitive load (e.g., Gennari *et al.*, 2018; Mattys *et al.*, 2014). Poorer performance under divided attention is generally attributed to competition for limited processing resources (Kahneman, 1973) and rapid attentional switches between the competing tasks (Moray, 1986). A correlate of the attentional-switch assumption is that, as attention is intermittently diverted to the competing task, input samples (or pulses) of the stimuli in the main task can be missed (Block and Zakay, 1996; Zakay and Block, 1995), leading to an under-estimation of the duration of those stimuli. We refer to this premise as the "pulse-skipping" hypothesis.

There is ample evidence that perceived duration is reduced under divided attention (e.g., see Taatgen *et al.*, 2007, for a review), but most studies have used stimulus durations in the range of several seconds. Few have considered durations within a range relevant to speech perception and, in particular, phoneme identification (i.e., <200 ms). For exception, Casini *et al.* (2009) asked participants to report which of two near-homophonous consonant-vowel-consonant (CVC) French words they heard under focused attention (single-tasking) versus divided attention (dual-tasking). The duration of the vowel was manipulated (from 150 to 310 ms). In the divided-attention condition, a red or green light-emitting diode was lit up in front of the participant during the play-back of the spoken word. Participants had to decide if the diode was red or green as fast as possible before deciding which word they heard. Divided attention led to a bias toward reporting the word whose final consonant was cued by shorter vowel duration (French word *cache* [ʃ] rather than *cage* [ʒ]). Thus, there is some evidence that the pulse-skipping hypothesis might apply to speech perception.

The goal of this study was to further explore the findings of Casini *et al.* (2009), focusing on a temporal cue known for its contribution to the voiced-unvoiced contrast in word-initial plosives in aspiration languages, namely, voice onset time (VOT; see, e.g., Abramson and Lisker, 1970). Studies by Gordon *et al.* (1993) and Mattys and Wiget (2011) provide mixed evidence for the effect of divided attention on VOT perception. Gordon *et al.* (1993) found that performing an arithmetic task during phonetic categorization led to slightly more frequent reports of long-VOT phonemes, contrary to expectation. However, because both VOT and f_0 cues were

co-varied in their experiment, it is difficult to evaluate the effect of divided attention on VOT alone. In the study of Mattys and Wiget (2011), performing a visual search task did not significantly affect categorization of the initial phoneme of a /gɪ/-/kɪ/ continuum, but discrimination of between-category contiguous steps on the same continuum was poorer.

In the present study, we used Mattys and Wiget's (2011) /gɪ/-/kɪ/ continuum and had listeners perform a phonetic categorization task and a discrimination task. Both tasks were administered under focused attention and divided attention. In the focused-attention condition, participants did the auditory tasks without any distraction. In the divided-attention condition, participants performed the auditory tasks while also performing a visual line-judgment task. With respect to phonetic categorization, we predicted that, should divided attention lead to a reduction in duration perception, as expected by the pulse-skipping hypothesis, there should be a shift in the categorization function toward reporting more /g/ (short VOT) under divided attention than under focused attention. In contrast, should divided attention merely decrease reporting accuracy without eliciting a directional bias, the categorization function should simply flatten under divided attention.

In the discrimination task, participants performed an AX task, with A as a fixed reference. In one condition, A was the shortest VOT endpoint (/gɪ/) and X was either that syllable (identical acoustically) or any of the longer VOT syllables (different acoustically, whether it was in the same or a different phonemic category) up to endpoint /kɪ/. In the other condition, A was the longest VOT endpoint (/kɪ/) and X was either that syllable (identical) or any of the shorter VOT syllables (different) up to endpoint /gɪ/. Under divided attention, the visual task was imposed during the playback of syllable X only. Attention was never divided during syllable A. On the assumption that the perceived duration of X is reduced under divided attention, divided attention should impair discrimination, but only when A is the short-VOT syllable /gɪ/. In this condition, the shorter perceived duration of X would reduce the durational contrast between A and X, and hence, make discrimination between those two syllables more difficult. Critically, however, divided attention should improve discrimination when A is the long-VOT syllable /kɪ/, because the shorter perceived duration of X would magnify the durational contrast between A and X, thus making discrimination easier. However, should divided attention simply decrease reporting accuracy without eliciting a directional bias, discrimination should decrease in both /gɪ/ and /kɪ/ reference conditions.

2. Method

2.1 Participants

Twenty-four York-based young adults participated in the study (21 female, $M_{\mathrm{age}} = 19.58$ years, 18–22 years). All were native speakers of British English. All participants were assessed to have normal hearing using threshold pure tone audiometry in accordance with the 2011 British Society of Audiology recommended procedure at 0.5, 1, 2, and 4 KHz. None of the participants exceeded a threshold of 20 dB hearing level at any of the frequencies in either ear. All participants reported normal or corrected-to-normal vision. The study was approved by the University of York Departmental Ethics Committee (identification number 2018–712).

2.2 Materials

The auditory materials consisted of 15 syllables varying in their VOT along a /gɪ/–/kɪ/ (nonword–nonword) continuum. These syllables were modified versions of the syllable /kɪ/ in Mattys and Wiget (2011), which was extracted from a token of the word *kiss* produced by a female Southern British English speaker and recorded at a sample rate of 44 kHz (sample resolution of 16 bits). The continuum was generated by shortening the aspiration of the /kɪ/ syllable in fifteen 3-ms steps from 56 ms for the /kɪ/ endpoint to 14 ms for the /gɪ/ endpoint. Thus, including the /ɪ/ vowel, which was always 86 ms, syllable duration ranged from 100 ms (/gɪ/) to 142 ms (/kɪ/).

The visual stimuli used in the concurrent visual task were modified versions of load stimuli used by Macdonald and Lavie (2011). They consisted of a black cross displayed against a white background. Depending on the trial, the cross had either a longer vertical arm or a longer horizontal arm (3.6° versus 3.8° visual angle).

2.3 Design and procedure

The experiment took place in two sessions separated by at least one day. In the first session, participants completed a phonetic categorization task. In the second session, they completed a phonetic discrimination task. Both tasks were performed under focused and divided attention. Thus, all test conditions were administered within subjects. Participants were tested in a sound-attenuated booth. They were seated in front of a 22-in. Dell E2209Wf (REV A00) computer monitor with a 59/60 Hertz refresh rate. The viewing distance was approximately 60 cm.

Auditory stimuli were delivered over Denon DN-HP700 headphones (Denon, Kawasaki, Japan) at 65 dB sound pressure level.

Phonetic categorization task. To assess phonetic categorization, a two-alternative forced-choice task was used. Participants heard each step of the /ɡɪ/–/kɪ/ continuum 4 times, for a total of 60 trials.

On each trial of the focused-attention condition, a participant heard a syllable followed by 1000 ms of pink noise. During the pink noise, a gray square (subtending 19.8° × 19.8°) was displayed against a white background. A prompt reading "G………K" was then displayed. Participants had up to 8000 ms to press one of two keys to indicate if they heard /g/ or /k/. The respective location of the two keys corresponded to the left–right position of the G and K letters on the monitor. The position was counterbalanced between participants. Upon key press, or at the end of the 8000-ms period, there was a 1000-ms inter-trial interval until the next syllable.

In the divided-attention condition, the procedure was the same, except that a cross was displayed on the monitor during the playback of the test syllable. Cross display time was 150 ms in all trials to keep load level constant for all syllables. After responding /ɡɪ/ or /kɪ/, another prompt appeared showing a horizontal line on one side of the monitor and a vertical line on the other. Participants had up to 8000 ms to decide which of the horizontal or vertical arms of the cross seen during the auditory syllable was longer using two keys whose location corresponded to the left–right position of the horizontal and vertical lines on the monitor. These keys were the same as those used for the phonetic categorization task. This was followed by a 1000-ms inter-trial interval before the next syllable was played alongside another cross. No feedback was provided for either task.

Participants completed the focused- and divided-attention conditions in a single session. Trials were distributed across four blocks (30 trials each), two under focused attention and two under divided attention. Focused- and divided-attention blocks alternated. The alternation order was counterbalanced across participants. Before the experiment, participants practiced the visual task (20 trials) without auditory input. No feedback was provided.

Phonetic discrimination task. To assess phonetic discrimination, an AX task was used. Pairs of syllables were played with an inter-stimulus interval of 500 ms. In half the pairs, A was the /ɡɪ/ endpoint (14 ms VOT) and X was either the same syllable (identical trials) or one of the other 14 syllables (different trials). In the other pairs, A was the /kɪ/ endpoint (56 ms VOT) and X was either that same syllable or any of the other 14 syllables. For each of the two sets, the "identical" pair was played 112 times and each of the 14 "different" pairs was played 8 times, amounting to 112 "different" trials.

In the focused-attention condition, each AX pair was followed by 1000 ms of pink noise, as in the categorization task. The prompt "same………different" was then displayed on the computer monitor until the participant made a response. Participants were given up to 8000 ms to press one of two keys corresponding to the respective locations of the words identical (left) and different (right) on the monitor. Upon key press, or at the end of the 8000-ms period, there was a 1000-ms inter-trial interval until the next AX pair.

In the divided-attention condition, the procedure was the same, except that a cross was displayed on the monitor during the playback of the X syllable of each AX pair. As in the categorization task, the cross was displayed for 150 ms. After participants responded whether the syllables were the same or different, a prompt for the visual task appeared. The rest of the trial was the same as in the categorization task.

Participants completed the focused- and divided-attention conditions in a single session. They were instructed to discriminate A and X based on acoustic, not merely categorical, differences. Trials were distributed across four blocks (112 trials each), two under focused attention and two under divided attention. The difference between the two blocks within an attention condition was that the A syllable of the AX pairs was the /ɡɪ/ endpoint in one of them and the /kɪ/ endpoint in the other. Focused and divided attention blocks alternated. The alternation order was counterbalanced across participants.

3. Results

3.1 Phonetic categorization

Categorization responses are shown in Fig. 1. A logistic curve was fitted to the data of each participant under focused and divided attention. To test whether divided attention led to an increase in /ɡɪ/ responses, we compared the point of subjective equality on the logistic curve (PSE, or 50% threshold) between the two conditions. The PSE was not significantly different, with VOT = 27 ms under focused attention and 26 ms under divided attention, $t(23) = 1.19$, $p = 0.244$. Likewise, the percentage of /kɪ/ responses, averaged across the 15 steps of the VOT continuum,
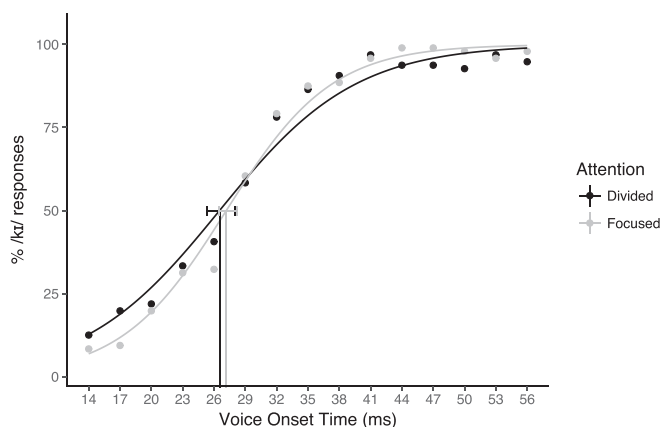
Fig. 1. Average responses and psychometric functions for the categorization of the /ɡɪ/-/kɪ/ continuum under divided and focused attention. The vertical lines show the PSE (or 50% threshold) and standard error of the mean for each of the two conditions.

showed no difference between the two conditions, both 67%, $t(23) = -0.36$, $p = 0.718$. Thus, divided attention did not increase the likelihood of /ɡɪ/ responses.

To test the possibility that divided attention, instead, flattened the categorization curve, a sign of increased response imprecision, we compared the slope of the logistic function for each participant under focused and divided attention. There was some evidence that the slope was shallower under divided (0.65) than focused (1.09) attention, $t(23) = 1.91$, $p = 0.069$. This effect appeared more strongly when we calculated the slope using the inter-quartile range method (IQR, e.g., Chauhan *et al.*, 1993; Strasburger, 2001), which minimizes the undue influence of outliers at the endpoints of the distribution, which was particularly notable at high VOT values. The IQR analysis consisted of measuring the distance on the VOT scale between the value corresponding to 25% and 75% /kɪ/ responses on the psychometric function, with larger values indexing a flatter slope. The IQR was 10 ms under focused attention and 15 ms under divided attention, $t(23) = 3.15$, $p = 0.004$, which confirms that the categorization function was flatter under divided than focused attention.

To further assess response imprecision, we compared the standard deviation (SD) across participants for each step of the continuum under focused and divided attention. The SD was larger under divided than focused attention, $t(14) = 2.22$, $p = 0.044$, suggesting that there was more variability among participants in the former than the latter.

Average performance on the visual task was 76% correct (chance = 50%), which suggests that the task was sufficiently challenging to produce a sizable attentional load without being so hard as to lead to disengagement from the task. Across participants, performance on the visual task did not correlate significantly with the threshold or the slope in the divided attention condition or with the difference in threshold or slope between the focused and divided attention conditions (all ps > 0.09). Thus, there was no evidence of trade-off between the auditory and the visual tasks.

3.2 Phonetic discrimination

AX discrimination accuracy is shown in Figs. 2(a) (/ɡɪ/ reference syllable) and 2(b) (/kɪ/ reference syllable). Because the hypothesis about a reduction in perceived duration under divided attention
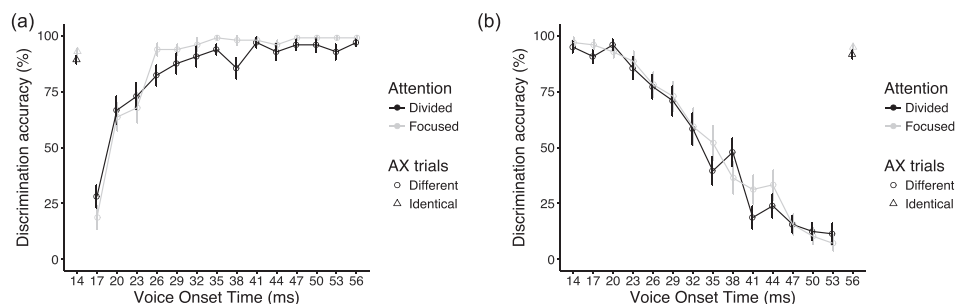

Fig. 2. Discrimination accuracy (% correct) in the AX task under divided and focused attention when the reference syllable is /ɡɪ/, which has a 14-ms VOT (a) and /kɪ/, which has a 56-ms VOT (b). Performance on the identical trials (triangles) is shown as disconnected from performance on the different trials (circles).

concerns primarily the different AX trials, accuracy was analyzed separately for the identical trials and the different trials. With respect to the identical trials, an analysis of variance (ANOVA) on percentage correct as a function of Attention (focused versus divided) and Reference Step (/ɡɪ/ versus /kɪ/) indicated a main effect of Attention, $F(1, 23) = 10.33$, $p = 0.004$, $\eta p^2 = 0.310$, with higher accuracy under focused than divided attention ($M = 94\%$, SD $= 7.3\%$, and $M = 90\%$, SD $= 9.3\%$, respectively). Neither the reference step, $F(1, 23) = 1.08$, $p = 0.310$, nor the interaction term, $F(1, 23) < 1$, was significant. Thus, participants were better at noticing that two syllables were identical under focused than divided attention, and this was independent of whether that syllable was /ɡɪ/ or /kɪ/.

With respect to the different trials, performance was averaged across all 14 different steps and submitted to an ANOVA by Attention (focused versus divided) and Reference Step (/ɡɪ/ versus /kɪ/). There was a main effect of Reference Step, $F(1, 23) = 96.319$, $p < 0.001$, $\eta p^2 = 0.807$, with higher accuracy when the reference was /ɡɪ/ ($M = 86\%$, SD $= 9.8\%$) than to /kɪ/ ($M = 54\%$, SD $= 15.3\%$). This difference probably reflects the generally larger just-noticeable difference at the high end than the low end of a scale (long VOT compared to short VOT), as formalized in Weber's law. Attention had a marginal effect on accuracy, $F(1, 23) = 4.18$, $p = 0.053$, with poorer accuracy under divided than focused attention. Critically, the interaction term was not significant, $F(1, 23) < 1$, which suggests that, contrary to the pulse-skipping hypothesis, the effect of Attention was comparable when the reference step was /ɡɪ/ or /kɪ/.

Average performance on the visual task, averaged across identical and different trials, was 77% correct. Across participants, performance on the visual task did not correlate significantly with any of the auditory discrimination measures (all $ps > 0.08$), indicating that there was no significant trade-off between the auditory and the visual tasks.

Taken together, the categorization and discrimination results fail to confirm the pulse-skipping hypothesis. Divided attention did not lead to a shift in phonetic categorization toward short VOT values and it did not lead to improved duration discrimination when the reference syllable in the AX task was /kɪ/. Instead, both categorization and discrimination data showed poorer consistency and accuracy under divided attention.

4. Discussion

In this experiment, we asked whether the hypothesis that time is perceived as shorter under divided attention (e.g., Block and Zakay, 1996; Zakay and Block, 1995) applies to the perception of speech sounds. Consistent with this hypothesis, referred to here as pulse-skipping, Casini et al. (2009) found that French vowel perception is shortened under divided attention. In contrast, related studies have shown mixed results (Gordon et al., 1993; Mattys and Wiget, 2011). However, because those studies were not set up to test the pulse-skipping hypothesis directly, we designed an experiment that was both a conceptual replication of the experiment of Casini et al. (2009), and a test of further implications of that hypothesis.

Contrary to expectations, listeners asked to categorize stop consonants varying in VOT, from /ɡɪ/ (short VOT) to /kɪ/ (long VOT), did not report more /ɡɪ/ responses under divided attention. Thus, the phonetic category boundary was unaffected by divided attention. However, the slope of the categorization function was shallower, indicating that divided attention increased imprecision around the category boundary. The expectation that shorter VOT perception under divided attention should improve discrimination in specific test conditions was not supported by the data either. Instead, divided attention impaired discrimination. In sum, this study shows that divided attention increases imprecision during speech sound perception, but it does not shorten perceived duration.

Our failure to replicate the results of Casini et al. (2009) with a VOT continuum could be attributed to several methodological differences between the two studies. First, Casini et al. (2009) used the duration of the vowel in CVC stimuli as a cue to the voicing of the following, final consonant (Fischer and Ohde, 1990; Raphael, 1972). Thus, duration was only an indirect cue to the critical contrast. In our study, duration (VOT) was a direct cue to voicing. A consequence of this methodological difference is that the duration range over which the effect of divided attention was measured was much shorter in our study (42 ms) than in the study of Casini et al. (2009) (160 ms). If the duration of the hypothetical "pulses" missed during attentional switching under divided attention exceeded our VOT temporal range, but not the range of Casini et al. (2009) it could partly explain why our manipulation only increased randomness in responding and not duration perception *per se*. Pulse duration has not been studied so far.

Second, because our durational variable, VOT, is generally described as being perceived categorically (e.g., Liberman et al., 1957), it might be more resilient to disruptions involving time perception than would be vowel perception, insofar as vowel duration has a continuous, rather than categorical, effect on post-vocalic voicing (Reinisch and Penney, 2019). Note, however, that

our VOT data did not exhibit a marked categorical profile in either the categorization or the discrimination condition. Conversely, the categorization responses of Casini *et al.* (2009) were more binary than would be expected from a strictly continuous model. Thus, the difference between our results and the results of Casini *et al.* (2009) is probably independent of the underlying perceptual mode between vowels versus consonants.

Third, differences in the tasks used as visual distraction could have contributed to the contrasted outcomes. In the study of Casini *et al.* (2009), participants had to categorize the color of a diode (red vs green) that lit up from the beginning to the end of the auditory syllable, thus varying with the duration of the syllable. In our study, the crosses used as visual stimuli were displayed for a fixed 150 ms and the co-occurring syllable was always shorter (from 100 to 142 ms). While the studies were comparable in that they both used a distractor task involving low-level vision, differences in the visual-to-auditory duration ratio could have affected the results. Research on low-level audio-visual integration suggests that relative timing between visual and auditory stimuli can influence duration perception, although the bulk of that research has focused on the effect of audition on vision rather than the other way around (e.g., Chen and Yeh, 2009). However, that research has also shown that small differences in duration between concurring auditory and visual stimuli led to poorer discrimination (Romei *et al.*, 2011), which could be one of the mechanisms at play here.

In sum, while time shrinkage under divided attention, as predicted by the pulse-skipping hypothesis, has been observed in a wide variety of tasks, its application to short durations relevant to phoneme perception is limited. Our results suggest that, with durations within a range of 50 ms, divided attention increases randomness in responding, as indicated by a shallower categorization curve and poorer discrimination performance, but not time perception itself. This pattern resembles, on a smaller scale, the flattening of categorization curves in cochlear-implant users (Winn and Litovsky, 2015) and in noisy conditions (Hazan *et al.*, 2009). It is also consistent with recent findings by Chiu *et al.* (2019) showing a substantial elevation in a just-noticeable-difference for the duration of synthetic speech sounds under divided attention.

References and links

Abramson, A., and Lisker, L. S. (**1970**). "Discriminability along the voicing continuum: Cross-language tests," in *Proceedings of the Sixth International Congress of Phonetic Sciences*, Academia, Prague.

Block, R. A., and Zakay, D. (**1996**). "Models of psychological time revisited," Time Mind **33**, 171–195.

Casini, L., Burle, B., and Nguyen, N. (**2009**). "Speech perception engages a general timer: Evidence from a divided attention word identification task," Cognition **112**, 318–322.

Chauhan, B. C., Tompkins, J. D., LeBlanc, R. P., and McCormick, T. A. (**1993**). "Characteristics of frequency-of-seeing curves in normal subjects, patients with suspected glaucoma, and patients with glaucoma," Invest. Ophthamol. Visual Sci. **34**, 3534–3540.

Chen, K. M., and Yeh, S. L. (**2009**). "Asymmetric cross-modal effects in time perception," Acta Psychologica **130**, 225–234.

Chiu, F., Rakusen, L. L., and Mattys, S. L. (**2019**). "Cognitive load elevates discrimination thresholds of duration, intensity, and $f0$ for a synthesized vowel," J. Acoust. Soc. Am. **146**, 1077–1084.

Fischer, R. M., and Ohde, R. N. (**1990**). "Spectral and duration properties of front vowels as cues to final stop-consonant voicing," J. Acoust. Soc. Am. **88**, 1250–1259.

Gennari, S. P., Millman, R. E., Hymers, M., and Mattys, S. L. (**2018**). "Anterior paracingulate and cingulate cortex mediates the effects of cognitive load on speech sound discrimination," NeuroImage **178**, 735–743.

Gordon, P. C., Eberhardt, J. L., and Rueckl, J. G. (**1993**). "Attentional modulation of the phonetic significance of acoustic cues," Cognit. Psychol. **25**, 1–42.

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., and Shakespeare, B. (**2009**). "Speech perception abilities of adults with dyslexia: Is there any evidence for a true deficit?," J. Speech, Lang. Hear. Res. **52**, 1510–1529.

Kahneman, D. (**1973**). *Attention and Effort* (Prentice-Hall, Englewood Cliffs, NJ).

Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (**1957**). "The discrimination of speech sounds within and across phoneme boundaries," J. Exp. Psychol. **54**, 358–368.

Macdonald, J. S., and Lavie, N. (**2011**). "Visual perceptual load induces inattentional deafness," Attn., Percept., Psychophys. **73**, 1780–1789.

Mattys, S., Barden, K., and Samuel, A. G. (**2014**). "Extrinsic cognitive load impairs low-level speech perception," Psychonomic Bull. Rev. **21**(3), 748–754.

Mattys, S. L., and Wiget, L. (**2011**). "Effects of cognitive load on speech recognition," J. Memory Lang. **65**, 145–160.

Moray, N. (**1986**). "Monitoring behavior and supervisory control," in *Handbook of Perception and Human Performance*, edited by K. R. Boff, L. Kaufman, and J. P. Thomas (Wiley, New York), pp. 1–51.

Navon, D., and Gopher, D. (**1979**). "On the economy of the human-processing system," Psychol. Rev. **86**, 214–255.

Raphael, L. J. (**1972**). "Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English," J. Acoust. Soc. Am. **51**, 1296–1303.

Reinisch, E., and Penney, J. (**2019**). "The role of vowel length and glottalization in German learners' perception of the English coda stop voicing contrast," Lab. Phonol. **10**, 18.

Romei, V., De Haas, B., Mok, R. M., and Driver, J. (**2011**). "Auditory stimulus timing influences perceived duration of co-occurring visual stimuli," Frontiers Psychol. **2**, 215.

J. Acoust. Soc. Am. **147** (6), June 2020

Chiu *et al.*   EL489

Strasburger, H. (**2001**). "Converting between measures of slope of the psychometric function," Percept. Psychophys. **63**, 1348–1355.

Taatgen, N. A., Van Rijn, H., and Anderson, J. (**2007**). "An integrated theory of prospective time interval estimation: The role of cognition, attention, and learning," Psychol. Rev. **114**, 577–598.

Winn, M. B., and Litovsky, R. Y. (**2015**). "Using speech sounds to test functional spectral resolution in listeners with cochlear implants," J. Acoust. Soc. Am. **137**, 1430–1442.

Zakay, D., and Block, R. A. (**1995**). "An attentional-gate model of prospective time estimation," in *Time and the Dynamic Control of Behavior*, IPA symposium, Liège, Belgium, pp. 167–178.