

Effects of cognitive load on speech recognition

Sven L. Mattys & Lukas Wiget

University of Bristol

Keywords: spoken-word recognition; speech perception; cognitive load; divided attention; Ganong

Corresponding author:
Sven Mattys
School of Experimental Psychology
University of Bristol
12a Priory Road
Bristol BS8 1TU
United Kingdom
Email: sven.mattys@bris.ac.uk

Abstract

The effect of cognitive load (CL) on speech recognition has received little attention despite the prevalence of CL in everyday life, e.g., dual-tasking. To assess the effect of CL on the interaction between lexically-mediated and acoustically-mediated processes, we measured the magnitude of the "Ganong effect" (i.e., lexical bias on phoneme identification) under CL and no CL. CL consisted of a concurrent visual search task. Experiment 1 showed an increased Ganong effect under CL. A time-course analysis of this pattern (Experiments 2 and 3) revealed that the Ganong effect decreased over time under optimal conditions, but it did not under CL. Thus, CL appears to be delaying (and perhaps preventing) listeners' ability to rely on fine phonetic detail to perform the sub-lexical task. This finding, along with an absence of measurable effects at the post-lexical level (Experiment 4) or at the lexical level (Experiment 5) and a clear negative effect of CL on perceptual discrimination (Experiment 6), suggests that the increased reliance on lexically-mediated processes under CL is the cascaded effect of impoverished encoding of the sensory input. Ways of implementing a link between CL and sensory analysis into existing models of speech recognition are proposed.

Models of speech recognition have traditionally focused on the mechanisms underlying the mapping between sensory input and lexical representations, with an emphasis on the relative contribution of sub-lexical cues and lexical-semantic knowledge. Insight has also been gained by studying the impact of adverse listening conditions on those mechanisms. Commonly studied adverse conditions include signal degradation (e.g., noisy, accented, or disordered speech), incomplete linguistic knowledge (e.g., non-native listeners), and impaired receptive abilities (e.g., listeners with cochlear implants). For instance, it has been shown that, while listeners rely more strongly on lexical-semantic knowledge than sub-lexical cues in optimal listening conditions, their relative reliance on some sub-lexical cues, rhythm and stress in particular, often increases when the sensory input is degraded by noise or non-canonical pronunciation (e.g., Bradlow & Alexander, 2007; Liss, Spitzer, Caviness, Adler, & Edwards, 1998; Mattys, White, & Melhorn, 2005; Smith, Cutler, Butterfield, & Nimmo-Smith, 1989).

However, adverse conditions resulting from a cognitive load have received less attention¹, even though such conditions are prevalent in everyday life, e.g., recognizing speech while performing another task. Here, we define cognitive load (CL) as any load whose effect on speech recognition arises not from an energetic distortion of the signal but from the recruitment of central processing resources due to concurrent attentional or mnemonic processing. At the core is the assumption that divided attention or short-term memory overload taps into processing resources needed for speech processing, and that such resources are in limited supplies (Kahneman, 1973). Dual-task paradigms can therefore distinguish those components of the speech system that are particularly dependent upon central processing resources from those that are less so (Navon & Gopher, 1979; Pashler, 1994). Thus, understanding how CL affects speech recognition is important because it can provide an insight into the interface between the speech

system and general cognition, a link often acknowledged in the speech-recognition literature but rarely implemented at a computational level (but see Grossberg, 1986; Mirman, McClelland, Holt, & Magnuson, 2008, for notable attempts).

The few studies that have addressed the issue of CL have shown that speech recognition can be substantially impaired by a secondary task, e.g., a visual search or mental arithmetic (Casini, Burle, & Nguyen, 2009; Gordon, Eberhardt, & Rueckl, 1993; Rönnerberg, Rudner, Foo, & Lunner, 2008; Toro, Sinnett, & Soto-Faraco, 2005; for an exception, see, Choi, Lotto, Lewis, Hoover, & Stelmachowicz, 2008). Critically, however, CL seems to cause more than just a generalized decrease in performance. For example, Wurm and Samuel (1997) showed that a tone-modulation monitoring task was more detrimental to a concurrent phoneme monitoring task when the latter was performed on true nonwords (e.g., *gizbozety*) than on word-like nonwords (e.g., *epilepty*). This result suggests that more processing resources are required for spoken stimuli with low than high levels of lexical involvement. Likewise, Fernandes, Kolinsky, and Ventura (2010) found that the extraction of novel words from a continuous speech stream is more greatly affected by a concurrent visual task when the word boundaries in the stream are cued by statistical regularities than by coarticulatory cues. The authors concluded that CL has a particularly detrimental effect on domain-general processes, e.g., statistical computation, and less effect on phonetic processing.

In the attention literature, CL involving divided attention has been shown to be particularly deleterious to the processing of stimuli that require active control, e.g., ignoring voice characteristics (Nusbaum & Schwab, 1986), or stimuli whose acoustic quality leads to more than one interpretation, e.g., a stimulus in between *sunny* and *funny* (Cowan, 1997; Francis & Nusbaum, 2009; Nusbaum & Magnuson, 1997; Nusbaum & Pisoni, 1985). In contrast, processes

involving automatic memory retrieval are less resource-demanding (e.g., Jacoby, 1991; Logan, 1988). Furthermore, work by Lavie and colleagues (e.g., Lavie, 2005, 2010) indicates that an important dissociation must be made between tasks involving a high degree of cognitive load, e.g., working memory demands, and tasks involving a high degree of perceptual load, e.g., physical complexity of the stimulus to process. High perceptual load increases participants' ability to focus on the main task and ignore irrelevant distractors whereas high cognitive load has the opposite effect. Greater distraction under cognitive load is believed to be due to a reduction of selective control and the inability to prevent attentional capture by salient, though irrelevant information (Lavie, Hirst, de Fockert, & Viding, 2004). Thus, performance on a main task is affected differently by perceptual and cognitive loads (see Mattys, Brooks, and Cooke, 2009, for a similar dissociation between energetic and informational masking) and cognitive load itself can be more or less disruptive depending on the nature of the main task.

Our own research reveals a systematic tendency for listeners under CL to downplay the contribution of acoustic detail and increase reliance on lexical-semantic knowledge (Mattys et al., 2009; Mattys, Carroll, Li, & Chan, 2010). In these two studies, listeners were asked to decide which of two words, e.g., *mild* or *mile*, they heard in a phrase like /mɑɪldɒpʃən/, the pronunciation of which ranged from /mɑɪld#ɒpʃən/ to /mɑɪl#dɒpʃən/, with # denoting acoustic word-boundary cues such as glottalisation, aspiration, lengthening. Under CL, listeners more frequently reported the word leading to the lexically acceptable segmentation solution (e.g., *mild*, leading to *mild option*) than the word leading to the lexically unacceptable segmentation solution (e.g., *mile*, leading to *mile doption*). This pattern, which we henceforth refer to as a CL-induced lexical drift, was observed under various types of CL: divided attention in dichotic

listening, concurrent rehearsal of words or syllables in working memory, concurrent visual search.

The exact nature of this CL-induced lexical drift is unclear. Since the listeners in the Mattys et al. (2009, 2010) studies gave their responses on a single rating scale (e.g., *mild* to *mile*, in eleven steps), the tendency to give more lexically-acceptable responses under CL could be due to either increased reliance on lexical-semantic knowledge (lexical locus) or reduced reliance on sub-lexical cues (sub-lexical locus). CL effects at a lexical locus could be implemented by any of the following, depending on the underlying model's architecture: increased activation of lexical nodes, decreased lexical activation thresholds, or increased lateral inhibition between lexical nodes (e.g., McClelland & Elman, 1986; Mirman et al., 2008; Norris, McQueen, & Cutler, 2000). CL effects at a sub-lexical locus could be implemented by decreasing lateral inhibition between sub-lexical nodes, reducing signal-to-noise ratio in the encoding of acoustic cues, or decreasing the perceptual sampling rate of acoustic detail (e.g., Gordon et al., 1993; McClelland & Elman, 1986). CL effects could also occur outside the lexical-access loop, with greater reliance on lexically viable responses under CL arising from a post-lexical, strategic bias for communicatively meaningful percepts (e.g., the "merging" stage in Norris et al.'s, 2000, Merge model). Neither our research (Mattys et al., 2009, 2010) nor the studies reviewed earlier or associated literature (e.g., Cutler, Mehler, Norris, & Segui, 1987; Eimas, Marcovitz Hornstein, & Payton, 1990; Eimas & Nygaard, 1992) provide a clear answer to these questions.

In the present study, we begin to explore the above alternatives by investigating the impact of CL on the so-called Ganong effect (Ganong, 1980). The Ganong effect refers to the influence of lexical knowledge on phoneme identification. For example, the phoneme identification curve along a /g/-/k/ continuum—where the voice onset time (VOT) of a syllable-onset stop consonant

is manipulated in small increments—usually shows an inflation of /k/ responses when the phoneme is followed by *-iss* (*giss-kiss* continuum, with the word underlined) and an inflation of /g/ responses when the phoneme is followed by *-ift* (*gift-kift* continuum). We chose this paradigm because it highlights the interplay between phonetic processing and lexical knowledge in a task that requires that listeners actively *suppress* lexical activation. Thus, the case for a lexical drift under CL will be more compelling if it can be observed when reliance on lexical information is explicitly discouraged. In addition, Ganong stimuli can be submitted to a variety of tests with no change to the stimuli themselves: phoneme identification with and without lexical context (e.g., *giss to kiss, gift to kift, gi to ki*), explicit lexical judgment (e.g., "is *?ift* an acceptable exemplar of the word *gift*?"), cross-modal semantic priming (e.g., does *?ift* facilitate the recognition of *birthday*?), AX phoneme discrimination (e.g., *gift₁* vs. *gift₂*). Thus, the versatility of Ganong stimuli makes the paradigm well suited to exploring the locus of CL on speech recognition.

Experiment 1 examines the standard Ganong effect with and without CL, as well as its time-course. Experiments 2 and 3 further investigate the time-course of CL in a deadline and a delay versions of Experiment 1, respectively. Using a lexical acceptability task, Experiment 4 tests a post-lexical account of the effect of CL on the Ganong effect. Using a cross-modal semantic priming paradigm, Experiment 5 tests a lexical account. Experiment 6 evaluates the impact of CL on low-level sensory analysis in an AX discrimination task. In all experiments, CL was created by having participants perform a concurrent visual search task in an array of colored shapes (see Figure 1). We used this task because of its non-linguistic nature and non-auditory modality; hence, any effect on the speech task is unlikely to result from simple modality-/domain-specific interference, and more likely to result from depletion of central processing resources (Styles, 1997). This task successfully elicited a lexical drift in Mattys et al. (2010).

Experiment 1

In this experiment, we tested a large number of participants (111) on a standard Ganong phoneme-identification task. The sample size allowed us not only to increase the stability of our patterns but also to run detailed post-hoc time-course analyses. Participants were assigned to either a baseline condition with no cognitive load (No-CL, similar to Ganong's [1980] Experiment 1) or to the same task under cognitive load (CL).

Method

Participants.

One hundred and eleven native speakers of English received course credit or a small honorarium for taking part in the experiment. None of them reported a history of speech or hearing difficulties. These criteria applied to all participants in this article. The participants were randomly assigned to the No-CL (N = 52) or CL condition (N = 59).

Materials.

The stimuli consisted of three 8-step continua: *giss-kiss* (nonword-word), *gift-kift* (word-nonword), and *gi-ki* (nonword-nonword). These stimuli were chosen because they have repeatedly been shown to yield a reliable Ganong effect. To create the continua, we asked a female native speaker of Standard Southern British English to utter multiple renditions of the words *gift* and *kiss*. The clearest token of each word was selected and split into the initial consonant (/g/ or /k/), the vowel (/ɪ/), and the coda (/ft/ or /s/). The /ɪ/ we kept originated from one of the *kiss* tokens because it was particularly clear and showed fairly neutral coarticulation. Using the two initial consonants, we created an 8-step VOT continuum from /g/ to /k/, with VOT values of 15 ms, 23 ms, 28 ms, 33 ms, 38 ms, 43 ms, 48 ms, and 56 ms. The eight steps of the continuum were then recombined with the /ɪ/ and the coda /ft/ or /s/ to form the *giss-kiss*, *gift-*

kift, and *gi-ki* continua. The average duration of the three continua was, respectively, 464 ms, 535 ms, and 121 ms.

The visual arrays used in the CL condition consisted of grids made of six rows and six columns. The grids were 9 cm by 9 cm. The 36 items in each grid were black squares and red triangles. These were arranged randomly in the grid (see an example in Figure 1A). However, half the grids contained a red square, which was the target participants were required to detect. The red square could be anywhere in the grid (Figure 1B).

Procedure.

Stimuli were recorded in a sound-attenuated booth and stored directly onto disk with a sample rate of 44 kHz and sample resolution of 16 bit, and delivered at 44 kHz 16 bit over Sennheiser HD 25 SP headphones at approximately 65 dB SPL. In both the No-CL and the CL conditions, each of the 24 syllables (8 *giss-kiss*, 8 *gift-kift*, and 8 *gi-ki*) was played 3 times, for a total of 72 trials delivered in a different random order for each participant. On each trial, participants were asked to decide if the first sound of the syllable was /g/ or /k/. They were explicitly encouraged to focus on the initial sound and not on the meaning of the syllable. Participants in the CL condition were additionally asked to pay attention to the array displayed on a 19-inch computer monitor in front of them during the playback of the syllable and search for a red square. The distance between the participant's head and the monitor was approximately 55 cm.

Timing within each trial was as follows. In the No-CL condition, an auditory syllable was immediately followed by the written message "G..... initial sound.....K" displayed on a computer monitor for 1 s. Participants had up to 10 s from the onset of the syllable to press one of two keys on a computer keyboard for /g/ or /k/. The respective location of the two keys on the

keyboard corresponded to the left-right position on the monitor. After key press, or at the end of the 10-s period, there was a 2-s inter-trial-interval. The next syllable was then played. The procedure for the CL condition was similar, except that a visual array was displayed on the computer monitor during the playback of the auditory syllable. To keep the level of CL constant for all syllables, the array was displayed for 560 ms in all trials. It was immediately followed by the phoneme identification message ("G... initial sound...K"). After key press, or at the end of a 10-s period, a second written message appeared, reading: "no...red square?...yes". Participants had up to 10 s from the onset of that message to press one of two keys on the computer keyboard. These keys were the same as those for the identification task. After key press, or at the end of a 10-s period, there was a 2-s inter-trial-interval. The next syllable was then played, along with the next visual array. Target-absent and target-present arrays were randomly assigned to the syllables.

Results and Discussion

The phoneme identification curves for the No-CL and CL conditions are displayed in Figures 2A-B. The Ganong effect, calculated as the area between the *_iss* and *_ift* curves averaged across the eight steps of the continuum, was clearly visible in both No-CL and CL. Critically, the Ganong effect was greater under CL than No-CL (33% vs. 25%). An analysis of variance performed on the participants' average proportions of /k/ responses across Load (No-CL vs. CL) and Context (*_iss* vs. *_ift*) showed that the Context effect (i.e., standard Ganong effect), $F(1, 110) = 258.24, p < .001$, interacted with Load, $F(1, 110) = 25.12, p < .001$. Thus, the Ganong effect was significantly larger under CL than No-CL.² There was no main effect of Load, $F(1, 110) = 3.00, p = .09$. Note that, while the *_iss* context appeared to create a weaker lexical bias than the *_ift* context, both contexts significantly departed from the *_i* context, $F(1,$

110) = 14.88, $p < .001$ and $F(1, 110) = 312.35$, $p < .001$, respectively. As for the responses in the _i context alone, CL led to a bias towards /k/ reports, $F(1, 110) = 6.46$, $p = .01$, but this bias was not significantly affected by Step (1 to 8), $F(7, 770) = 1.85$, $p = .08$ (Figure 2C). The /k/ bias under CL is difficult to explain, but it could reflect either an enhancement of right-hand (dominant) responses or a preference for consonants' unmarked form (i.e., unvoiced) when the system is under stress.

With respect to the visual-search task itself (CL condition only), performance did not significantly differ between the _iss (80%), _ift (78%), and _i (80%) contexts, $F(2, 220) = 1.15$, $p = .22$, but variations were noted across the eight steps, $F(7, 770) = 17.35$, $p < .001$ (73-85%), and Context and Step interacted, $F(14, 1540) = 3.71$, $p < .001$. Upon visual inspection (Figure 2D), this interaction did not reveal any clear or meaningful patterns. In particular, there was no indication that performance was poorer for mid-continuum tokens.

In sum, the increased Ganong effect under CL shows that the lexical drift reported by Mattys et al. (2009; 2010) is robust and arises even with a task that explicitly involves suppressing lexical access. Before exploring whether this lexical drift resulted from a post-lexical bias, from genuinely enhanced lexical activation (lexical locus), or from a lexical bias due to impoverished sensory analysis (sub-lexical locus), we explored a crucial difference between the No-CL and CL conditions, namely, the time it took participants to give their phoneme-identification responses.

Time-course considerations

Participants' phoneme-identification responses were substantially slower in the CL than No-CL condition (average: 1487 ms vs. 918 ms, respectively, measured from syllable onset. Median: 1368 ms vs. 817 ms. See Figure 3, black lines). Could the lexical drift under CL be a

by-product of delayed responses rather than a direct consequence of divided attention? Previous studies have shown that the magnitude of the Ganong effect can be substantially affected by latency. The direction of this interaction is unclear, however. For example, Pitt and Samuel (1993, 2006) and McQueen (1991) have reported an attenuation of the Ganong effect with slow RTs, whereas Burton, Baum, and Blumstein (1989), Fox (1984), and Miller and Dexter (1988) have reported an attenuation of the Ganong effect with fast RTs. Although these discrepancies could be due to differences in design, phoneme identity/position, and stimulus length, the sensitivity of the Ganong effect to latency suggests that RTs should be considered in our interpretation.

We fit linear regression models to the raw identification responses (/k/ vs. /g/) in each of the four conditions of the Context (_iss vs. _ift) by Load (No-CL vs. CL) design along the time domain (Figure 4). To capture the most stable parts of the distributions, responses produced under 450 ms or in the slowest five percentiles of the relevant condition were not entered in the analyses (No-CL: > 1869ms; CL: > 2614 ms). Generalized linear mixed-effect models were then built, in which Load (No-CL vs. CL), Context (_iss vs. _ift), and RT (numeric) were used as predictors of the listeners' proportion of /k/ responses. Using log-likelihood χ^2 tests to compare model fits (e.g., Baayen, Davidson, & Bates, 2008), we found a Context-by-RT interaction in both the No-CL and the CL conditions, $\chi^2(1) = 62.73, p < .001$, and $\chi^2(1) = 11.10, p < .001$, respectively, reflecting an attenuation of the Ganong effect over time in both cases. Critically, however, as indicated by the significant improvement of the model when a 3-way interaction term involving Load (No-CL vs. CL) was added, $\chi^2(1) = 39.40, p < .001$, this attenuation was more gradual in the CL than the No-CL condition. Furthermore, an analysis of variance of participants' individual regression intercepts in a Context (_iss vs. _ift) by Load (No-CL vs. CL)

model revealed an effect of Context, $F(1, 109) = 96.63, p < .001$, no effect of Load, $F(1, 109) < 1$, and, importantly, no interaction, $F(1, 109) = 2.58, p = .11$. The lack of interaction suggests that the size of the Ganong effect at short RTs was not affected by Load.

These RT analyses lead to two conclusions. First, although robust at fast RTs, the Ganong effect tends to decrease with slow RTs, in line with Pitt and Samuel's (1993, 2006) and McQueen's (1991) findings. We argue that time allows listeners to suppress the automatic activation of lexical representations and more accurately consider the acoustic memory trace left by the input. Second, what we call a lexical drift under CL might be better described as a delayed capacity to suppress lexical activation. Thus, CL might not so much operate by *increasing* reliance on lexical knowledge as by *prolonging* it and delaying the contribution of acoustic detail to the listener's percept.

A drawback with the above analyses is that the patterns at fast and slow RTs were extrapolated from relatively small data sets (or, at least, from data sets unevenly distributed across the No-CL and CL conditions). In the next experiments, we explored these time-specific patterns more directly by either encouraging participants to respond by a deadline (Experiment 2) or by forcing them to delay their responses (Experiment 3). It should be noted that deadline and delay methods come with their own task demands, which could possibly affect the magnitude of the lexical bias and/or the effect of CL. For example, time pressure has been shown to modulate lexical processing via a decrease of activation thresholds (e.g., McElree, 1996) or, conversely, an increase in input gain (e.g., Kello, 2004). Likewise, delayed-response procedures could introduce memory demands interacting with CL's own demands. Thus, we do not present Experiments 2 and 3 as exact cross-sections of the time-course analysis in Experiment 1, but,

rather, as a way of checking that the time-course pattern in Experiment 1 was not simply due to unrepresentative tail-end data.

Experiment 2

Participants were presented with randomly intermixed No-CL and CL trials. Fast responses were encouraged by overtly timing out responses produced later than 1500 ms after syllable onset. Our goal was to align the distribution of RTs in the CL condition with that in the No-CL condition as much as possible. If CL affects speech processing by genuinely increasing lexical activation, the lexical drift under CL should be visible even at short RTs. However, if CL affects speech processing by simply prolonging lexical activation, the lexical drift should be minimal, if not nil, at short RTs, as suggested in Experiment 1 (cf. Figure 4).

Method

Participants, materials, and procedure.

This experiment included 35 participants. The materials were those of Experiment 1. The 72 No-CL trials and 72 CL trials from Experiment 1 were presented to all participants in a fully randomized order. The procedure was similar to that of Experiment 1, except that participants only had 1500ms to produce their phoneme-identification response. If they did not meet this deadline, the next trial started (in the No-CL condition) or the prompt for the visual search task appeared (in the CL condition). Participants were explicitly instructed to perform the task as fast as possible.

Results and Discussion

Ninety percent of the trials were responded to before the deadline. Not surprisingly, the deadline was met less often in the CL than No-CL condition, 86% vs. 94%, $F(1, 34) = 44.39$, $p < .001$. Although the average RTs under CL still lagged behind those under No-CL (average: 1045

ms vs. 900 ms. Median: 1037 ms vs. 876 ms), the overlap between the No-CL and CL RT distributions was clearly better than that in Experiment 1 (Figure 3, dark grey lines). Figures 5A-B show the identification curves for the No-CL and CL conditions. An analysis of variance, with Context (_iss vs. _ift) and Load (No-CL vs. CL) as within-subject factors, showed an effect of Context, $F(1, 34) = 33.79, p < .001$, no effect of Load, $F(1, 34) < 1$, and critically, no interaction between Context and Load, $F(1, 34) = 1.56, p = .22$. Thus, the size of the Ganong effect did not differ between the CL and No-CL conditions at fast RTs. An analysis of variance restricted to the four mid-continuum steps (steps 3, 4, 5, 6) did not reveal a significant interaction either, $F(1, 34) = 2.05, p = .16$. Evidence for an increased Ganong effect under CL was only visible when the analysis was restricted to the two mid-most steps (steps 4 and 5), $F(1, 34) = 6.86, p = .01$. As a whole, however, there was little evidence that CL caused an increased Ganong effect when a deadline was imposed on the participants' responses—unless the stimuli were acoustically extremely underspecified. This result is consistent with the time-course analysis in Experiment 1 and suggests that CL does not have much effect on speech processing at short latencies.

As for responses in the _i context, the No-CL and CL conditions did not differ from each other, $F(1, 34) < 1$, and Load did not interact with Step, $F(7, 238) < 1$ (Figure 5C). With respect to the visual-search task (CL condition only), performance differed across the _iss, _ift, and _i contexts, $F(2, 68) = 16.28, p < .001$, with better performance in the _iss (79%) and _ift (77%) contexts than in the _i context (70%), $F(1, 34) = 28.24, p < .001$, and $F(1, 34) = 19.48, p < .001$, respectively. The _iss and _ift contexts did not differ from each other, $F(1, 34) = 2.34, p = .14$. This pattern could indicate that lexically-mediated responses (only possible in the _iss and _ift conditions) require fewer processing resources, hence allowing more resources to be available for the visual-search task. Variations were also noted across the eight steps, $F(7, 238) = 3.09, p <$

.005 (range: 68% to 82%), but Context and Step did not interact, $F(14, 476) = 1.25, p = .23$. As in Experiment 1, visual inspection of the responses did not reveal any clear or meaningful patterns (Figure 5D).

Experiment 3

This experiment was conceptually similar to Experiment 2, except that participants were not allowed to respond to the phoneme identification task until a prompt was displayed.

Method

Participants, materials, and procedure.

This experiment, which included 34 participants, was identical to Experiment 2, except that a delay was imposed on participants' responses in the phoneme identification task. To do so, we included 1500 ms of idle time between the offset of the syllable and the phoneme identification prompt ("G...initial sound...K"). The monitor was blank during the idle time and key presses during that period were not recorded. As in Experiment 1, once the prompt appeared, participants had up to 10 s to respond. After key press, or at the end of the 10-s period, the next trial started (in the No-CL condition) or the prompt for the visual-search task appeared (in the CL condition). Participants were explicitly told that they should not rush to perform the phoneme-identification because responses given before the prompt would not be recorded.

Results and Discussion

As can be seen in Figure 3 (light grey lines), the RT distributions in the CL and No-CL conditions were comparable (average: 2125 ms vs. 2103 ms., respectively Median: 1957 ms vs. 1978 ms, respectively). Figures 6A-B show the identification curves for the No-CL and CL conditions. An analysis of variance, with Context (_iss vs. _ift) and Load (No-CL vs. CL) as within-subject factors, showed a main effect of Context, $F(1, 33) = 36.97, p < .001$, no effect of

Load, $F(1, 33) < 1$, and an interaction between Context and Load, $F(1, 33) = 13.53, p = .001$.

Thus, as predicted by the time-course pattern in Experiment 1, at slow RTs, the Ganong effect was greater under CL than No-CL. Note that the present experiment showed an asymmetrical Ganong effect between the *_iss* and *_ift* contexts (relative to the *_i* context), with a strong effect in the *_ift* context, $F(1, 33) = 67.06, p < .001$, and no effect in the *_iss* context, $F(1, 33) < 1$. This asymmetry, which was visible in Experiments 1-2 as well, could be due to a greater capacity for *gift* than *kiss* to induce a lexical response (for lexical or acoustic reasons) or to an intrinsic bias for /k/ responses in the *_i* context. Because categorization asymmetry is frequently observed in phoneme-identification studies and because it does not undermine the main conclusion in the present study, we did not investigate it further.

As for the responses in the *_i* context itself, we found no effect of Load, $F(1, 33) = 1.07, p = .31$, or interaction between Load and Step, $F(7, 231) = 1.23, p = .29$ (Figure 6C). With respect to the visual-search task (CL condition only), performance did not differ significantly between the *_iss* (80%), *_ift* (78%), and *_i* (81%) contexts, $F(2, 66) = 2.90, p = .06$, but variations were found by Steps, $F(7, 231) = 5.69, p < .001$ (range: 72% to 87%), and Context interacted with Step, $F(14, 462) = 2.35, p < .005$. As before, visual inspection of the visual-search responses did not reveal any interpretable patterns (Figure 6D).

Taken together, Experiments 1-3 suggest that inhibiting lexical activation during phoneme identification is more difficult when the task is performed under divided attention (CL) than focused attention (No-CL). Why is it so hard to inhibit lexical information under CL? In the following experiments, we examined three possibilities. First, CL could merely encourage a response bias whereby reference to lexical knowledge is strategically favored when processing resources are limited. Lexical and sub-lexical processing would not be affected per se. Second,

the effect could be explained at a lexical level if CL acts on the speech system by, for example, decreasing lexical activation thresholds or increasing activation gain (see Mirman et al., 2008, for additional options). In this case, we would expect CL to increase lexical reliance even when the opportunity to respond strategically is minimized. Third, the lexical effect under CL could simply be a cascaded consequence of impoverished encoding of the input signal, without intrinsic changes at the lexical or post-lexical levels. If so, discrimination of low-level aspects of the signal should be impoverished under CL in both lexically meaningful and meaningless stimuli. Experiments 4 to 6 constitute a first step towards distinguishing these alternatives.

Experiment 4

Experiment 4 explored a post-lexical explanation of the CL-induced lexical drift by asking whether listeners favored a lexical interpretation of the signal under CL compared to No-CL when performing a meta-linguistic version of the task in Experiment 1. Participants were asked to explicitly judge whether or not the test syllables of Experiments 1-3 were acceptable exemplars of the word *kiss* or *gift*. The syllables from the *_i* context were excluded from this experiment. We reasoned that a post-lexical locus of CL should manifest itself by more liberal lexical-acceptability judgments, that is, a propensity to report more words under CL than No-CL.

Method

Participants, materials, and procedure.

This experiment included 38 participants. The materials were those of Experiment 1, except that the syllables on the *gi-ki* continuum were removed. The remaining 48 No-CL trials and 48 CL trials from Experiment 1 were presented to all participants in a fully randomized order. On each trial, participants were asked to decide whether or not the syllable they heard was an acceptable exemplar of the word *kiss* or *gift*. They were not asked to discriminate between the

two words, as the syllables on the *giss-kiss* continuum obviously called for a judgment relative to *kiss* and those on the *gift-kift* continuum called for a judgment relative to *gift*. The timing of events within a trial was the same as in Experiment 1, except that the prompt displayed on the computer monitor for the lexical-acceptability judgment had “good” on the left side of the monitor and “bad” on the right side.

Results and Discussion

The eight categories on the continuum were relabeled such that, in both the *_iss* and *_ift* conditions, Step 1 indicated a syllable with a lexically appropriate onset (word end, *kiss* or *gift*) and Step 8 indicated a syllable with a lexically inappropriate onset (nonword end, *giss* or *kift*). The data were collapsed across Context (*_iss* vs. *_ift*). Lexical-acceptability judgment curves for the No-CL and CL conditions are displayed in Figure 7. An analysis of variance performed on the participants' acceptability scores across Load (No-CL vs. CL) and Step (word to nonword) showed no main effect of Load, $F(1, 37) < 1$. Thus, contrary to the post-lexical hypothesis, listeners did not respond to CL by simply relaxing their lexical-acceptability criterion compared to No-CL. However, a Load by Step interaction, $F(7, 259) = 4.50, p < .001$, indicated that participants under CL were less discriminant about what constituted good vs. bad exemplars of words—the judgment curve was flatter under CL than No-CL. In fact, when responses given to Steps 5 to 8 were re-coded into their mirror values (1 into 0 and 0 into 1) so as to evaluate all steps on a single "judgment confidence" scale, a strong Load effect appeared $F(1, 37) = 16.20, p < .001$, with judgment confidence lower under CL than No-CL. Performance in the visual-search task (CL condition) ranged from 58% to 80% across steps, but with no indication of a linear trend, $F(1, 37) = 2.43, p = .13$, as shown in Figure 7.

In sum, these results suggest that the lexical drift under CL cannot be reduced to a simple post-lexical preference to report real words. Instead, listeners' discriminability between what they considered acceptable and unacceptable lexical exemplars was attenuated under CL at both the word and nonword ends of the continuum. We come back to this point in Experiment 6.

Experiment 5

In this experiment, we tested whether the CL-induced lexical drift found with the Ganong task was manifest when the lexical level itself was probed and opportunities to respond strategically were minimized. To do so, we used a cross-modal semantic priming task (Swinney, Onifer, Prather, & Hirshkowitz, 1979; Warren, 1972). In the critical trials, we measured the lexical-decision latencies to a visual target (e.g., *birthday*) when it was preceded either by one of our Ganong stimuli acting as a semantically associated prime (e.g., *gift*, *?ift*, *kift*) or by a semantically unrelated prime (e.g., *lake*). The prime was presented under CL or No-CL. If CL leads to an increase in lexical activity—for example by lowering lexical activation thresholds or increasing activation gain—cross-modal semantic priming should be greater under CL than No-CL. This lexical boost might be particularly visible with lexically ambiguous primes (e.g., *?ift*). For consistency with the other experiments, the test stimuli were only those used so far, that is, the *kiss-giss* and *gift-kift* Ganong series.

Method

Participants, materials, and design.

This experiment included 98 participants. The critical auditory primes were the two end-point syllables of each continuum (i.e., two word primes: *kiss* and *gift* and two nonword primes: *giss* and *kift*) and the two mid-most syllables on each continuum. We used this restricted set of syllables (8 instead of the 16 original ones) to minimize repetition of similar versions of the

primes. Forty-six additional words and 46 nonwords were recorded, all of them monosyllabic. These were filler primes, except for the words *home* and *lake*, which were used as baseline primes for the *kiss-giss* and *gift-kift* series, respectively. The word *home* was chosen because it was semantically unrelated to *lips*, the visual target for the *kiss-giss* primes and, likewise, *lake* was semantically unrelated to *birthday*, the target for the *gift-kift* primes. As visual targets, in addition to *lips* and *birthday*, we selected 38 words and 50 nonwords, with an approximately equal number of monosyllables and disyllables in both sets. There were more nonwords than words to compensate for the test targets, which were all words. The stimuli and task used as CL were those of the previous experiments (the 6-by-6 visual arrays).

Before counterbalancing, the experiment included 10 test trials to be presented under CL and under No-CL. In five of them, the target was *lips*, and the primes were: *kiss*, the two mid-most ambiguous *?iss* syllables (steps 4 and 5), *giss*, and *home*. In the other five, the target was *birthday*, and the primes were: *gift*, the two mid-most ambiguous *?ift* syllables (steps 4 and 5), *kift*, and *lake*. To minimize repetition, however, each participant was only presented with one of the two ambiguous primes (step 4 or 5) per set, and this was counterbalanced across participants. Furthermore, of the eight trials, four were presented in No-CL and four in CL, which was counterbalanced across participants as well.

To break the association between the test primes and the test targets, some of the test primes were also followed by a nonword target; likewise, some of the baseline primes were also followed by a semantically associated target word. In total, each participant was exposed to 100 trials (8 test and 92 fillers). Half the visual targets were words, half were nonwords. Word and nonwords targets could be preceded by a word or a nonword prime, in approximately equal proportions (the approximation is due to the ambiguous test primes being impossible to classify

as words or nonwords). When the prime and the target were both words, they were semantically associated in approximately half the trials. Half the trials were presented under CL and half under CL. The 100 trials were randomized for each participant. The experiment was preceded by eight practice trials made of materials unused in the main experiment.

Procedure.

On each No-CL trial, the participants heard a prime, followed by a 400-ms silent interval. A letter string was then displayed on a computer monitor for 1 s. The participant's task was to decide if the letter string was a word or a nonword as fast as possible, pressing the right shift key for words and the left shift key for nonwords. An inter-trial-interval of 2 s started either upon button press of after 4 s if no response was given. CL trials were similar, except that each CL trial started with the presentation of a fixation point in the center of the monitor for 500 ms followed by a 500-ms interval. The auditory prime and the visual array were then presented simultaneously. As in the No-CL trials, the target appeared 400 ms after the end of the prime and participants had 4 s to perform the lexical-decision task. Upon button press (or after 4 s, if no response was given), a message appeared, reading: "no...red square?...yes", prompting participants to report the presence or absence of a red square in the visual array. They had up to 10 s from the onset of that message to press one of two keys on the computer keyboard. These keys were the same as those for the lexical-decision task. After key press, or at the end of the 10-s period, there was a 2-s inter-trial-interval. Target-absent and target-present arrays were randomly assigned to the CL trials.

Results and Discussion

Lexical-decision latencies were measured from the onset of visual target presentation. Incorrect responses to word targets and correct responses two standard deviations from the mean

(computed on all 100 trials and separately for each participant) were discarded. Altogether, the discarded responses amounted to 2.3 % of the test trials (1.5 % incorrect). Mean lexical decision latencies, accuracy, and priming are reported in Table 1. Priming effects, plotted in Figure 8, were calculated as the difference between the latency to a target (e.g., *lips*) preceded by a related prime (e.g., *kiss*, *?iss*, or *giss*) and the latency to the same target preceded by the associated unrelated prime (e.g., *home*).

An analysis of variance, factoring Load (No-CL vs. CL) and Type of Prime (word [*kiss* and *gift*], ambiguous [*?iss* and *?ift*], nonword [*giss* and *kift*]), showed an effect of Type of Prime, $F(2, 192) = 4.87, p < .01$, no effect of Load, $F(1, 93) < 1$, and no interaction, $F(2, 181) < 1$. There was no Load effect for any of the three types of prime, all $F_s < 1$. The word primes produced more priming than the ambiguous and the nonword primes, $F(1, 96) = 4.20, p < .05$, and $F(1, 96) = 7.24, p < .01$, respectively. There was no significant priming difference between the ambiguous and nonword primes, $F(1, 96) = 2.14, p = .15$. The Type of Prime effect is broadly consistent with studies showing graded lexical activation as a function of phonetic goodness of fit (Andruski, Blumstein, & Burton, 1994; Connine, Blasko, & Titone, 1993). As for the visual-search task, performance did not differ across the four prime conditions (word, ambiguous, nonword, baseline), $F(3, 282) < 1$, and none of the pairwise comparisons reached .05 significance.

This experiment showed no evidence that lexical activation is enhanced under CL, which is all the more striking for a task in which greater reliance on lexical knowledge would lead to facilitation, a feature likely to be desirable for a system under strain. The lack of enhanced lexical effect was not due to a failure to attend to the prime under CL since the pattern and magnitude of priming were comparable under No-CL and CL. Thus, it is unlikely that the CL-

induced lexical drift found with the Ganong task and previously (Mattys et al., 2009, 2010) originates from a change in excitability of lexical representations. In the next experiment, we consider a sub-lexical locus for the effect of CL.

Experiment 6

Experiment 6 explores the possibility that CL impoverishes the analysis of the sensory input. Under this scenario, the CL-induced lexical drift would be a secondary consequence of poor sensory encoding rather than evidence of an alteration at the lexical or post-lexical level *per se*. In this experiment, listeners performed an AX discrimination task on pairs of syllables from Experiment 1. In line with previous studies (e.g., Abramson & Lisker, 1970), we used different levels of coarseness: Test syllables were either adjacent along the continuum ("one step," fine discrimination) or two steps away from each other ("two steps," coarser discrimination). If CL impoverishes sensory analysis, discrimination should be lower under CL than No-CL. This might be particularly noticeable for the mid-points of the continuum, where sensitivity to small changes has been reported to be most acute, at least with consonants (e.g., Liberman, Harris, Hoffman, & Griffith, 1957; for a recent review, see Holt & Lotto, 2010).

Method

Participants, materials, and procedure.

This experiment included 27 participants. The materials were those of Experiment 1, except that end-point syllables (15-ms and 56-ms VOT) were omitted because their distance from the adjacent syllables was larger (8 ms) than the distance between all other syllables (5 ms). Thus, there were six syllables for each of the *giss-kiss*, *gift-kift*, and *gi-ki* continua. Pairs of syllables were assembled in a context-specific fashion, that is, *_iss* syllables were only paired with other *_iss* syllables, *_ift* syllables with other *_ift* syllables, and *_i* syllables with other *_i*

syllables. Within each context, we created five pairs of adjacent syllables ("one step": 23-28ms, 28-33ms, 33-38ms, 38-43ms, 43-48ms) and four pairs of syllables two steps away from each other ("two steps": 23-33ms, 28-38ms, 33-43ms, 38-48ms). In addition to these nine "different" pairs, there were six "same" pairs, in which each syllable was paired with itself.

Each "different" pair (e.g., 23-28ms) led to a quadruplet of pairs: two "different" pairs in opposite orders (e.g., 23-28ms and 28-23ms) and two "same" pairs (e.g., 23-23ms and 28-28ms). Each context (_iss, _ift, _i) therefore included 36 pairs (9 x 4), for a total of 108 pairs. These 108 pairs were played under No-CL and CL, leading to 216 trials in all.

The 216 trials were presented to all participants in a fully randomized order. On each trial, participants were asked to decide whether or not the two syllables of a pair started with the same sound. The two syllables of a pair were separated by 500 ms. The timing of events within a trial was the same as in Experiment 1, except that the prompt displayed on the computer monitor for the discriminability task had "different" on the left side of the monitor and "same" on the right. Participants pressed one of two keys for "different" or "same." The respective location of the two keys on the keyboard corresponded to the left-right position on the monitor.

Results

Discrimination accuracy was calculated as the percentage of correctly responded-to pairs on a quadruplet-of-pairs basis (e.g., 23-28, 28-23, 23-23, 28-28), for each participant and each cell of the design. We preferred this method over d' and $Beta$ from Signal Detection Theory (Green & Swets, 1966), because of the large proportion of 0% and 100% values among the cells of the design—a consequence of using a single item—which are problematic for calculating statistics derived from z scores. We analyzed discrimination accuracy separately for the two Coarseness conditions (one-step vs. two-step). Results are shown in Figure 9. An analysis of

variance on the accuracy scores in the one-step condition, with Load (No-CL vs. CL), Context (*_i*, *_iss*, *_ift*), and Step (23-28, 28-33, 33-38, 38-43, 43-48) as within-subject factors, showed a main effect of Load, with poorer discrimination in the CL than No-CL condition, $F(1, 26) = 7.24, p = .01$. Performance was also significantly affected by Step, $F(4, 104) = 7.24, p < .001$, with better discrimination in the middle of the continuum than near the end points. An interaction between Load and Step, $F(4, 104) = 3.03, p = .02$, indicated that the Load effect only reached significance for two mid-continuum pairs: 33-38, $F(1, 26) = 6.55, p = .02$, and 38-43, $F(1, 26) = 5.91, p = .02$. None of the other effects or interactions reached significance.

A comparable analysis in the two-step condition (Steps: 23-33, 28-38, 33-43, 38-48), showed a main effect of Load, $F(1, 26) = 12.35, p = .002$, again suggesting that CL negatively affected discrimination. Likewise, an effect of Step, $F(3, 78) = 14.30, p < .001$, indicated better discrimination for pairs near the middle than the ends of the continuum. Additionally, discrimination differed by Context, $F(2, 52) = 7.26, p = .002$, with lower discrimination in the *_i* context than in the *_iss* and *_ift* contexts, $F(1, 26) = 13.00, p = .001$, and $F(1, 26) = 5.11, p < .03$, respectively. There was no difference between the *_iss* and *_ift* contexts, $F(1, 26) = 2.77, p = .11$. Finally, a Context by Step interaction, $F(6, 156) = 3.92, p < .005$, indicated that discrimination rose more sharply around the mid-continuum for the *_iss* and *_ift* contexts than for the *_i* context. None of the other interactions reached significance.

We also tested whether CL had an effect on listeners' bias for responding "same" or "different." Although the overall bias for responding "same" was notable (78% of all responses, which is not surprising, given the subtle distinction under study), Load did not significantly modulate this bias: One-step: $F(1, 26) = 2.44, p = .13$ (No-CL:82%; CL: 81%); Two-step: $F(1, 26) = 2.65, p = .12$ (No-CL: 72%; CL: 76%). Average performance on the visual-search task was

84 %, which is higher than in Experiments 1-5, probably on account of the longer exposure time to the visual display in this experiment. Scores were not meaningfully affected by Load, Condition, or Coarseness, all $ps > .05$.

Discussion

These results lead to several observations. First, discriminating stop consonants contrasted on small VOT differences was difficult but possible: Better-than-chance discrimination was particularly salient around the phoneme-category boundary, in line with a large body of evidence on categorical perception of consonants (e.g., Liberman et al., 1957; Repp, 1984). Discrimination was possible even when the two syllables of a pair differed by only 5 ms of VOT (one-step condition). Second, and more critically, discrimination of VOT differences was negatively affected by the presence of CL. This decrement was independent of whether or not lexical information was available in the signal—the CL effect was comparable in the *_iss*, *_ift*, and *_i* contexts. Third, the effect of CL on listeners' performance was restricted to discrimination. CL did not affect their overall bias to respond "same" or "different."

The latter two observations suggest that CL's effect on the speech system is visible as early as during the sensory analysis of the input. Ways in which impoverished signal encoding can lead to a lexical drift are proposed in the General Discussion.

General Discussion

This study investigated the effect of cognitive load (CL) on the relative contribution of lexical and sub-lexical information to speech recognition. Specifically, our goal was to elucidate the lexical drift that Mattys et al. (2009; 2010) reported when listeners' processing resources are taxed by divided attention or short-term memory overload. We hypothesized that listeners' relatively greater reliance on lexical information under CL could have a post-lexical, lexical, or a

sub-lexical locus. Establishing the locus of CL on the speech system is important not only because CL differs from other adverse conditions in theoretically critical ways (cf. Lavie's, 2005, distinction between perceptual and cognitive load), but also because the interaction between speech processing and CL can potentially shed light on the broader issue of language-cognition interface.

Using a phoneme-identification task (following Ganong, 1980), we found that listeners were more prone to base their identification responses on lexical information when phoneme identification was performed under CL than under No-CL, even though the task explicitly required that listeners suppress access to lexical representations (Experiment 1). This pattern not only replicates Mattys et al.'s (2009; 2010) lexical-drift phenomenon, it also shows that increased reliance on lexical information under CL cannot be reduced to a deliberate strategy to cope with strained processing resources (see also Experiment 4).

A time-course analysis of this effect (Experiments 2-3) revealed two important findings. First, overall, the size of the Ganong effect decreased over time. A similar attenuation was reported by Pitt and Samuel (1993, 2006) and McQueen (1991). This decrease of lexical interference suggests that, in the initial stages of processing, the activation of lexical representations is automatic and takes precedence over detailed phonetic analysis. Phonetic detail becomes more fully integrated once lexical influence can be controlled (and suppressed) at a later stage of processing. Thus, our data are incompatible with the claim that lexical information increases over time (e.g., Fox, 1984) due to confirmatory top-down feed-back (McClelland & Elman, 1986) or as compensation for the decay of phonetic echoic traces (e.g., Efron, 1970; Massaro, 1972). Second, and more critically, the larger Ganong effect under CL was due primarily to a lack of attenuation of the lexical effect at late RTs. That is, whereas

responses became less lexically-mediated over time in the No-CL condition, they did not in the CL condition. Thus, CL appears to interfere with detailed phonetic analysis. Impaired perceptual discrimination under CL (Experiment 6) is consistent with this scenario. Indeed, at an early stage of processing, lexical reliance in the No-CL and CL conditions would be comparable because, at this stage, lexically-based responses are dominant regardless of whether phonetic detail is accessible (No-CL) or not (CL). Impoverished sensory analysis under CL would only be noticeable at a later stage as listeners, no longer fully constrained by lexical information, would lack the adequate phonetic detail to do the task on an acoustic-phonetic basis.

With the above proposal in mind, two questions need to be answered: (1) How can a sensory-degradation account of CL lead to an increased Ganong effect, but not to a change in lexical activation (Experiment 5) or to a meta-linguistic preference for lexical responses (Experiment 4)? (2) What mechanisms underlie the effect of CL on sensory analysis?

We will consider the first question in the context of two contrasted architectures, one with feed-back between lexical and sub-lexical units (e.g., TRACE, by McClelland & Elman, 1986) and one without (e.g., Merge, by Norris et al., 2000). The two architectures have slightly different ways of dealing with tasks involving phoneme-level responses. In a feed-back architecture, phoneme responses can only be produced at a sub-lexical level (albeit influenced by lexical knowledge via top-down connections). In a feed-forward architecture, phoneme responses can be made either at a sub-lexical level or at a level where the contents of the sub-lexical and lexical levels have merged (the "decisional" level in Merge, which does not exist in feed-back architectures), but in most tasks involving a certain degree of listener control (e.g., phoneme categorization), responses are made at the decisional level.

In a feed-forward architecture, an increased Ganong effect under CL without manifestation at the lexical or post-lexical levels is expected if it can be shown that the impact of sensory degradation is not as severe on the lexical level as it is on the decisional level. If so, the lexical contribution to the decisional level would be broadly similar under No-CL and CL, whereas the contribution from the sub-lexical level would be weaker under CL. The net result of this imbalance would be an increased Ganong effect under CL. The different impact of CL at the lexical versus decisional levels could be due, on the one hand, to the lack of obvious lexical competition in the tasks we used (e.g., *lift* was never presented as an option to *gift*) and, on the other hand, to the fact that the information received by the decisional level is assumed to be a copy of the information at the sub-lexical level, with all its incompleteness or ambiguity. A similar scenario holds for feed-back models, where phoneme identification is read out directly from the sub-lexical level. Comparative simulations are needed to establish the extent to which CL-induced disruption of sensory analysis has cascaded consequences at the lexical level per se.

At this level of description at least, both feed-back and feed-forward architectures are equally compatible with our CL results. A key question for feed-forward models, however, is whether the results could be accounted for at a locus that does not directly involve sub-lexical sensory analysis. In particular, could our patterns originate from an effect of CL on the decisional level rather than the sub-lexical level? The discrimination data in Experiment 6 suggest that it is unlikely because of the low-level, perceptual nature of the task and the fact that reduced discrimination under CL was found for meaningful and meaningless syllables alike. However, the extent to which any task can be said to exclusively involve low-level perception has been a perennial debate in the literature (e.g., Magnuson, McMurray, Tanenhaus, & Aslin,

2003; McQueen, 2003) and is contingent on various design specifications (e.g., Pisoni & Tash, 1974).

It should be pointed out that our results and claim are at odds with Fernandes et al.'s (2010) finding that CL (in the form of a secondary visual task) had little effect on the use of coarticulatory cues for speech segmentation in artificial-language learning experiments. An important difference between their experiments and ours³, however, is that the acoustic cues in the Fernandes et al. study were far more salient and unambiguous than those in the present ones. This important contrast confirms that CL's effect on sensory analysis is restricted to *detailed* phonetic processing, perhaps those details for which there are no clear and categorical phonological consequences, e.g., small VOT differences (the present experiments) vs. meaningful allophonic distinctions (Fernandes et al.).

It is also useful to compare our results with Mirman et al.'s (2008) behavioral and computational data of the effect of attention on speech recognition. Mirman et al.'s work was based on the observation that lexical knowledge can facilitate phoneme monitoring (e.g., faster detection of /k/ in *logic* than in *hologic*; e.g., Rubin, Turvey, & Van Gelder, 1976), but that this facilitatory effect can be modulated by changing the proportion of word and non-word fillers in the experiment, with greater lexical facilitation when the fillers are mostly words than when they are mostly non-words (e.g., Eimas & Nygaard, 1992; Mirman et al., 2008). This modulatory effect is generally thought to be the consequence of a shift of attention between a sub-lexical locus (majority of non-word fillers) and a lexical locus (majority of word fillers). Mirman et al. showed that this shift could be simulated within the feed-back architecture of TRACE by altering the excitability of the lexical layer—resulting in an increase or decrease of feed-back to the phoneme layer. In particular, reduced lexical attention was successfully simulated by either

decreasing the lexical nodes' responsiveness to the phoneme nodes (wide and diffuse lexical activation) or adding a negative bias to the lexical nodes (lack of over-threshold lexical activation).

Mirman et al.'s (2008) study and ours differ in two important ways. First, Mirman et al. were specifically interested in simulating the effect of *attention* on lexical feed-back to sub-lexical processing. Our focus was on *cognitive load*, broadly construed. Although divided attention was an important feature of our CL manipulation, it was simply used as a way of decreasing resources for the speech task as a whole, not as a principled way of selectively orienting listeners' attention to or away from lexical representations. Second, Mirman et al. were mostly concerned about simulating a *decrease* of attention to lexical processing, whereas our goal was to understand why CL causes a *relative increase* in lexical reliance (cf. "lexical drift," Mattys et al., 2009, 2010). Given these two differences, the contrasted locus of explanation in Mirman et al. (lexical) and in this present study (sub-lexical) should not necessarily be seen as incompatible. For instance, it is conceivable that simulating an *increase* in attention to lexical representations could be achieved by decreasing the excitability of the phoneme nodes or the lateral inhibition between them. Likewise, using Mirman et al.'s procedure, it would be useful to test whether enhancing attention to lexical representations (implicitly or explicitly) and the presence of CL lead to comparable patterns of results. In other words, it is possible that increased attention to lexical representations and increased attention to sub-lexical representations involve mechanisms taking place in different layers of the speech system.

Our second question concerned the mechanism responsible for impoverished sensory analysis under CL. While the present results do not provide a direct answer to this question, the literature on the effect of attention on auditory and phonetic perception points to a few

possibilities. With respect to current theories of attention, our results are compatible with both Nusbaum and Pisoni's (1985) claim that central capacities are required when processing acoustically ambiguous stimuli and Lavie's (2005) claim that cognitive load impairs processes involving selective control. Indeed, in our experiments, CL was particularly detrimental to the processing of ambiguous tokens on the VOT continuum, where more than one categorical interpretation was possible. At the same time, there was no evidence that lexical activation *per se* was affected by CL, which is consistent with the view that processes involving retrieval from long-term memory are comparatively less resource-demanding (e.g., Jacoby, 1991; Logan, 1988).

From a computational point of view, impoverished sensory analysis could result from either a reduction in lateral inhibition between phoneme nodes or a reduction in the excitability/excitation of these nodes. In the former case, the capacity of phoneme nodes to be excited by the sensory input would be unaffected by CL, but their ability to decrease the activation of competing nodes would be dampened, which would particularly impair phoneme discrimination. In the case of a reduction of excitability of the phoneme nodes themselves, one possibility is that CL exerts its effect on the temporal computation of the sensory input. According to a domain-general timer hypothesis (e.g., Coull, Vidal, Nazarian, & Macar, 2004; Macar, Grondin, & Casini, 1994), the time estimation of time-bound events is based on the registration of the number of temporal pulses that make up an event. Divided attention would cause perceivers to miss pulses, and hence, lead them to under-estimate the duration of events. Indeed, Casini, Burle, and Nguyen (2009) found that listeners asked to identify which of two near-homophonous French CVC words they heard tended to report the one with the shorter vowel when the task was performed under divided attention (color-discrimination task). Thus, by

extension, a decrease in input processing sampling rate under CL would lead to a random loss of sensory information, which would be detrimental for duration-based distinctions in particular, (e.g., VOT, vowel length) and lead to weaker activation of phoneme nodes in general. An entirely time-based explanation does not seem satisfactory, however, since it predicts a general bias for perceiving /g/ (short VOT) over /k/ (long VOT) under CL, which was not found.

Another way of modeling CL's effect on phoneme activation is to treat CL as a regulator of signal-to-noise ratio. In this conceptualization, a key role of attention is to optimize the signal-to-noise ratio in the encoding of acoustic cues, with discriminant (strong) cues being given priority (Gordon, Eberhardt, & Rueckl, 1993). Under divided attention, system noise would fail to be filtered out, causing speech cues to be masked, with the magnitude of the masking effect proportional to the strength of the cue. For instance, while VOT has been shown to be a stronger cue to voicing distinction than the onset frequency of the fundamental frequency (onset F0) (Abramson & Lisker, 1985), Gordon et al. found that attentional demands—a concurrent arithmetic task—affected listeners' reliance on VOT more than on onset F0. This pattern could not be reduced to an account whereby time-based cues, such as VOT, are particularly sensitive to attentional demands (Casini et al., 2009), because a "beat" - "bit" discrimination task revealed that the vowel duration cue less affected by attentional demands than were the formant patterns. Gordon et al. successfully simulated these data by decreasing the signal-to-noise ratio of cues independently within a Signal Detection framework. Their account was also shown to be consistent with rare cases in which changes in cue weighing under divided attention can enhance performance rather than inhibit it (Choi et al., 2008).

In summary, we found that the lexical drift observed under CL (Mattys et al., 2009, 2010) is likely to be a cascaded consequence of impoverished sensory analysis (Experiment 6) rather

than a direct modification by CL of lexical activation or post-lexical biases (Experiments 4 and 5). The CL-induced lexical drift takes time to manifest itself, however, because phonetic details are not immediately available to the system even under optimal conditions (Experiments 1-3). Speech scientists have repeatedly stated that models of speech recognition need to have provision for attention-based phenomena (e.g., McClelland, Mirman, & Holt, 2006; various commentaries on Merge, 2000). The present study provides an account of how this can be achieved.

References

- Abramson, A., S. & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the Sixth International Congress of Phonetic Sciences* (pp. 569-573). Prague: Academia.
- Abramson, A., S. & Lisker, L. (1985). Relative power of cues: F0 shift versus voice timing. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 25-33). New York: Academic Press.
- Andruski, J. E., Blumstein, S. E., & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, *52*, 163-187.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.
- Bradlow, A. R. & Alexander, J. A. (2007). Semantic-contextual and acoustic-phonetic enhancements for English sentence-in-noise recognition by native and non-native listeners. *Journal of the Acoustical Society of America*, *121*, 2339-2349.
- Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 567-575.
- Casini, L., Burle, B., Nguyen, N. (2009). Speech perception engages a general timer: Evidence from a divided attention word identification task. *Cognition*, *112*, 318-322.
- Choi, S., Lotto, A., Lewis, D., Hoover, D., & Stelmachowicz, C. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language, and Hearing Research*, *51*, 1042-1054.
- Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*, 193-210.
- Coull, J. T, Vidal, F., Nazarian, B., & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, *303*, 1506-1508.
- Cowan, N. (1997). *Attention and memory: An integrated framework*. New York: Oxford University Press.
- Efron, R. (1970). The minimum duration of a perception. *Neuropsychologia*, *8*, 57-63.
- Eimas, P. D., Marcovitz Hornstein, S. B., & Payton, P. (1990). Attention and the role of dual codes in phoneme monitoring. *Journal of Memory and Language*, *29*, 160-180.
- Eimas, P. D. & Nygaard, L. C. (1992). Contextual coherence and attention in phoneme monitoring. *Journal of Memory and Language*, *31*, 375-95.
- Fernandes, T., Kolinsky, R., & Ventura, P. (2010). Cognitive noise is also noise: The impact of attention load on the use of statistical information and coarticulation as speech segmentation cues. *Attention, Perception, & Psychophysics*, *72*, 1522-1532.

- Fox, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 526-540.
- Francis, A. L. & Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psychophysics*, *71*, 1360-1374.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 110-125.
- Gordon, P. C., Eberhardt, J. L., & Rueckl, J. G. (1993). Attentional modulation of the phonetic significance of acoustic cues. *Cognitive Psychology*, *25*, 1-42.
- Green, D.M., Swets J.A. (1966) *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Grossberg, S. (1986). The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In E. C. Schwab and H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines, Vol 1: Speech perception* (pp. 187-294). Academic Press.
- Holt, L. L. & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception & Psychophysics*, *72*, 1218-1227.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kello, C. T. (2004). Control over the time-course of cognition in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 942-955.
- Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, *9*, 75-82.
- Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*, 143-148.
- Lavie N., Hirst A., De Fockert J.W., Viding E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339-354.
- Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*, 447-454.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358-368.
- Liss, J. M., Spitzer, S., Caviness, J.N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *Journal of the Acoustical Society of America*, *104*, 2457-2466.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.

- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*, 1-36.
- Macar, F., Grondin, S., & Casini, L. (1994). Controlled attention sharing influences time estimation. *Memory and Cognition, 22*, 673-686.
- Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science, 27*, 285-298.
- Massaro, D.W. (1972). Perceptual images, processing time, and perceptual units in auditory perception. *Psychological Review, 79*, 124-145.
- Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology, 59*, 203-243.
- Mattys, S. L., Carroll, L. M., Li, C. K. W., & Chan, S. L. Y. (2010). Effects of energetic and informational loads on speech segmentation by native and non-native speakers. *Speech Communication, 52*, 887-899.
- Mattys, S. L., White, L., & Melhorn, J. F (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General, 134*, 477-500.
- McClelland, J. L. & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.
- McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences, 10*, 363-369.
- McElree, B. (1996). Accessing short-term memory with semantic and phonological information: A time-course analysis. *Memory & Cognition, 24*, 173-187.
- McQueen, J.M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 433-443.
- McQueen, J. M. (2003). The ghost of Christmas future: Didn't Scrooge learn to be good? Commentary on Magnuson, McMurray, Tanenhaus, and Aslin (2003). *Cognitive Science, 27*, 795-799.
- Miller, J. L., & Dexter, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 369-378.
- Mirman, D., McClelland, J. L., Holt, L. L., & Magnuson, J. S. (2008). Effects of attention on the strength of lexical influences on speech perception: Behavioral experiments and computational mechanisms. *Cognitive Science, 32*, 398-417.
- Navon, D. (1984). Resources – A theoretical soup stone? *Psychological Review, 91*, 216-234.
- Navon, D. & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review, 86*, 214-255.

- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23, 299-370.
- Nusbaum, H. C. & Magnuson, J. S. (1997) Talker normalization: Phonetic constancy as a cognitive process. In K. A. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing*. Academic Press.
- Nusbaum, H. C. & Pisoni, D. B. (1985). Some constraints on the perception of synthetic speech. *Behavior Research Methods, Instruments, and Computers*, 17, 235-242.
- Nusbaum, H. C. & Schwab, E. X. (1986). The role of attention and active processing in speech perception. In E.C. Schwab & H.C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Vol. I. Speech perception* (pp. 113-157). San Diego: Academic Press.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116, 220-244.
- Pisoni, D. B. & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, 15, 285-290.
- Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 1-27.
- Pitt, M. A., & Samuel, A. G. (2006). Word Length and Lexical Activation: Longer is Better. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1120-1135.
- Repp, B. H. (1984). Categorical perception: Methods, issues, findings. In J. Lass (Ed.), *Speech and Language: Advances in Basic Research and Practice*, Vol. 10 (pp. 243-355). New York: Academic Press.
- Rönnerberg, J, Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, 47, S171-S177.
- Rönnerberg, J, Rudner, M., Lunner, T., & Zekveld, A. A. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise & Health*, 12, 263-269.
- Rubin, P., Turvey, M. T., Van Gelder, P. (1976). Initial phonemes are detected faster in spoken words than in spoken nonwords. *Perception & Psychophysics*, 19, 394-398.
- Smith, M. R., Cutler, A., Butterfield, S., & Nimmo-Smith, I. (1989). The perception of rhythm and word boundaries in noise-masked speech. *Journal of Speech and Hearing Research*, 32, 912-920.
- Styles, E. A. (1997). *The psychology of attention*. Bucks, UK: Taylor and Francis Psychology Press.
- Swinney, D., Onifer, W., Prather, P., & Hirshkowitz, M. (1979). Semantic facilitation across modalities in the processing of individual words and sentences. *Memory & Cognition*, 7, 159-165.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). The consequences of diverting attention within and across sensory modalities on statistical learning. *Cognition*, *97*, B25-B34.

Warren, R.E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology*, *94*, 90-100.

Wurm, L. H., and Samuel, A. G. (1997). Lexical inhibition and attentional allocation during speech perception: Evidence from phoneme monitoring. *Journal of Memory and Language*, *36*, 165-187.

Authors' note

Correspondence should be sent to Sven Mattys, School of Experimental Psychology, University of Bristol, 12A Priory Road, Bristol BS8 1TU, UK, E-mail: Sven.Mattys@bris.ac.uk. We thank Clive Frankish, Jeff Bowers, Nina Kazanina, Markus Damian, Katharine Barden, Katherine Wilkinson for their insightful comments and help with data collection. This study was made possible thanks to a grant from the Leverhulme Trust (F/00 182/BG) and from the ESRC (RES-062-23-2746) to S. L. Mattys.

Footnotes

¹ In this article, we do not consider cognitive load that is secondary to perceptual degradation or that arises from processing linguistic complexity. For recent reviews on these topics, see, e.g., Rönnerberg, Rudner, Lunner, and Zekveld (2010) and Lewis, Vasishth, and Van Dyke (2006).

² To assess the possibility that the larger Ganong effect under CL could have been the result of a radical change in strategy under dual-tasking conditions rather than the result of graded resource depletion due to cognitive load (Navon, 1984; Navon & Gopher, 1979), we ran a new version of Experiment 1 with two grades of CL (low-CL vs. high-CL) instead of No-CL vs. CL. The high-CL condition was the same as the CL condition in Experiment 1 (6-by-6 visual arrays). The low-CL condition used 3-by-3 visual arrays. The rest of the materials was the same as in Experiment 1. So was the procedure, except that low-CL and high-CL trials were randomly mixed. We tested 34 participants. An analysis of variance performed on Load (low-CL vs. high-CL) and Context (_iss vs. _ift) showed a main effect of Context, $F(1, 33) = 67.07, p < .001$, and no main effect of Load, $F(1, 33) < 1$. Critically, however, a significant Context-by-Load interaction, $F(1, 33) = 8.59, p = .006$, revealed that the Ganong effect was larger in the high-CL condition (29%) than in the low-CL condition (20%). Performance on the visual task was better under low-CL than high-CL (96% vs. 85%, respectively), $F(1, 33) = 118.33, p < .001$. Thus, listeners' increased reliance on lexical knowledge under CL was genuinely caused by reduced processing capacity rather than a radical strategic change in how the task was performed.

³ For a detailed comparison between Fernandes et al.'s (2010) position and ours, see Mattys et al. (2009, p. 235).

Table

Table 1. Latencies (% correct) and priming effects in Experiment 5. ***: $p < .001$; **: $p < .01$; *: $p < .05$; †: $p < .10$; ns: $p > .10$.

	No-CL		CL		Vis. Task (CL)
	RT (Acc.)	Prim.	RT (Acc.)	Prim.	RT (Acc.)
WD	528 (98)	69 ***	594 (98)	70 **	400 (74)
Amb.	560 (97)	37*	635 (98)	29 †	414 (74)
NW	588 (97)	5 ns	649 (97)	15 ns	478 (79)
Basl.	597 (99)		664 (96)		497 (78)

Figure Captions

Figure 1. Examples of visual displays used as cognitive load in all experiments. A. Target-absent display. B. Target-present display, with the target (red square) in the second column and fifth row.

Figure 2. Experiment 1. A-B. Percentage of /k/ responses for the *giss-kiss*, *gi-ki*, and *gift-kift* continua with no cognitive load (No-CL) and under cognitive load (CL). The average Ganong effect for the No-CL and CL conditions, shown as a percentage value, was calculated as the area between the *_iss* and *_ift* curves averaged across the eight steps of the continuum. C. Responses for the *gi-ki* continuum only. D. Accuracy in the visual-search task.

Figure 3. Probability density of phoneme-identification RTs for the No-CL and CL conditions in Experiments 1 (black lines), 2 (dark grey lines) and 3 (light grey lines). RTs are from the onset of the test syllables. The area under each curve equals 1.

Figure 4. Experiment 1. Scatterplot of the proportion of /k/ responses over time for the Context (*_iss* vs. *_ift*) by Load (No-CL vs. CL) design. Data points are responses averaged over 50-ms bins. Regression lines were fitted to the raw data. The bar chart displays the frequency distribution of data for the No-CL (solid grey) and CL (white with black border) conditions on a bin-by-bin basis. The scatterplot and the bar chart should be considered together when assessing the reliability of the data points along the RT dimension.

Figure 5. A-B. Experiment 2. Percentage of /k/ responses for the *giss-kiss*, *gi-ki*, and *gift-kift* continua without cognitive load (No-CL) and with cognitive load (CL), under a 1500-ms response deadline. C. Responses for the *gi-ki* continuum only. D. Accuracy in the visual-search task.

Figure 6. A-B. Experiment 3. Percentage of /k/ responses for the *giss-kiss*, *gi-ki*, and *gift-kift* continua without cognitive load (No-CL) and with cognitive load (CL), under a 1500-ms response delay. C. Responses for the *gi-ki* continuum only. D. Accuracy in the visual-search task.

Figure 7. Experiment 4. Percentage of syllables judged to be lexically acceptable for each step of the acoustic continuum from a clear word (*kiss* or *gift*) to a clear nonword (*giss* or *kift*), with no cognitive load (No-CL) and under cognitive load (CL). The data are collapsed across the *_iss* and *_ift* contexts. Performance on the visual-search task is shown as crosses (right Y axis).

Figure 8. Experiment 5. Average priming (ms) and error bars as a function of type of prime and CL.

Figure 9. Experiment 6. Percentage of correct discrimination between adjacent syllables on the /g/-/k/ continuum (one step, upper four figures) and syllables two steps away from each other (two steps, lower four figures). Results are broken down by context in the three leftmost figures, and collapsed across context in the rightmost figure.

Figure 1.

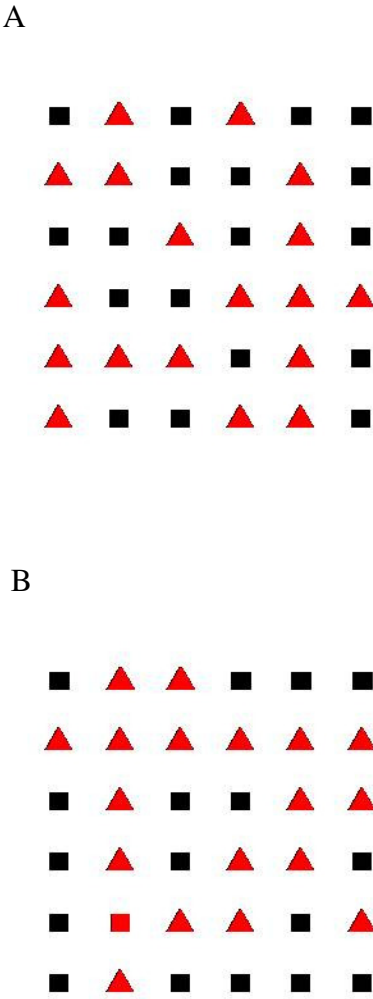
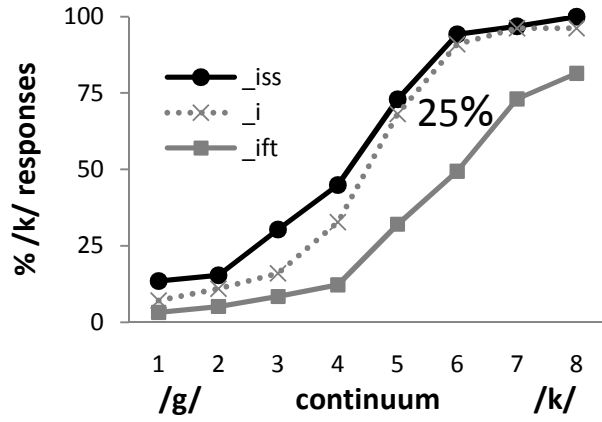
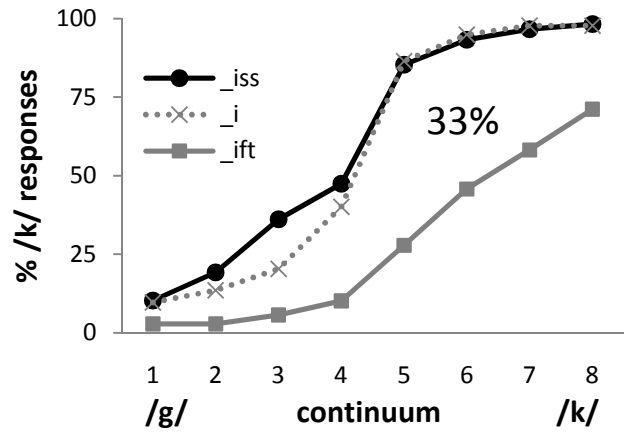


Figure 2.

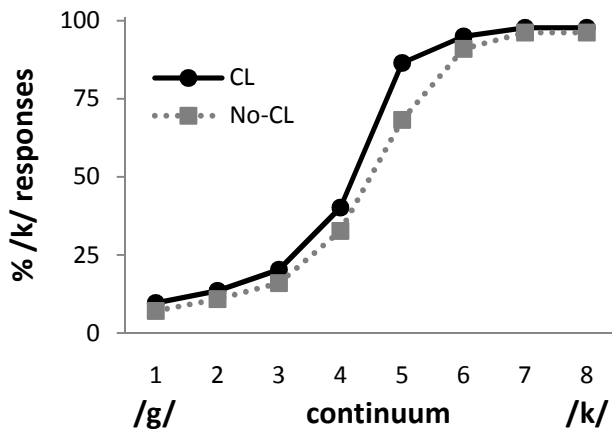
A. No-CL



B. CL



C. _i only



D. Visual Search Task

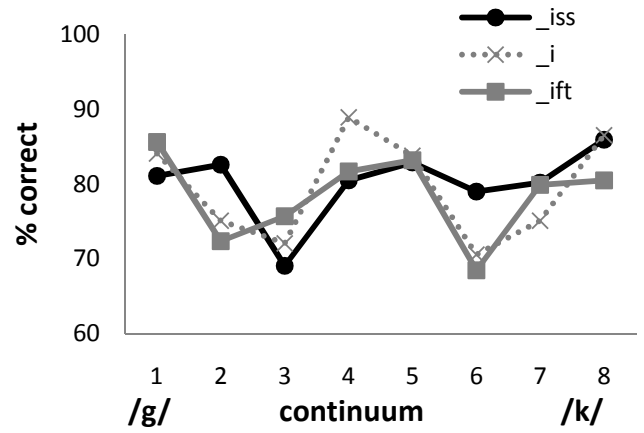


Figure 3.

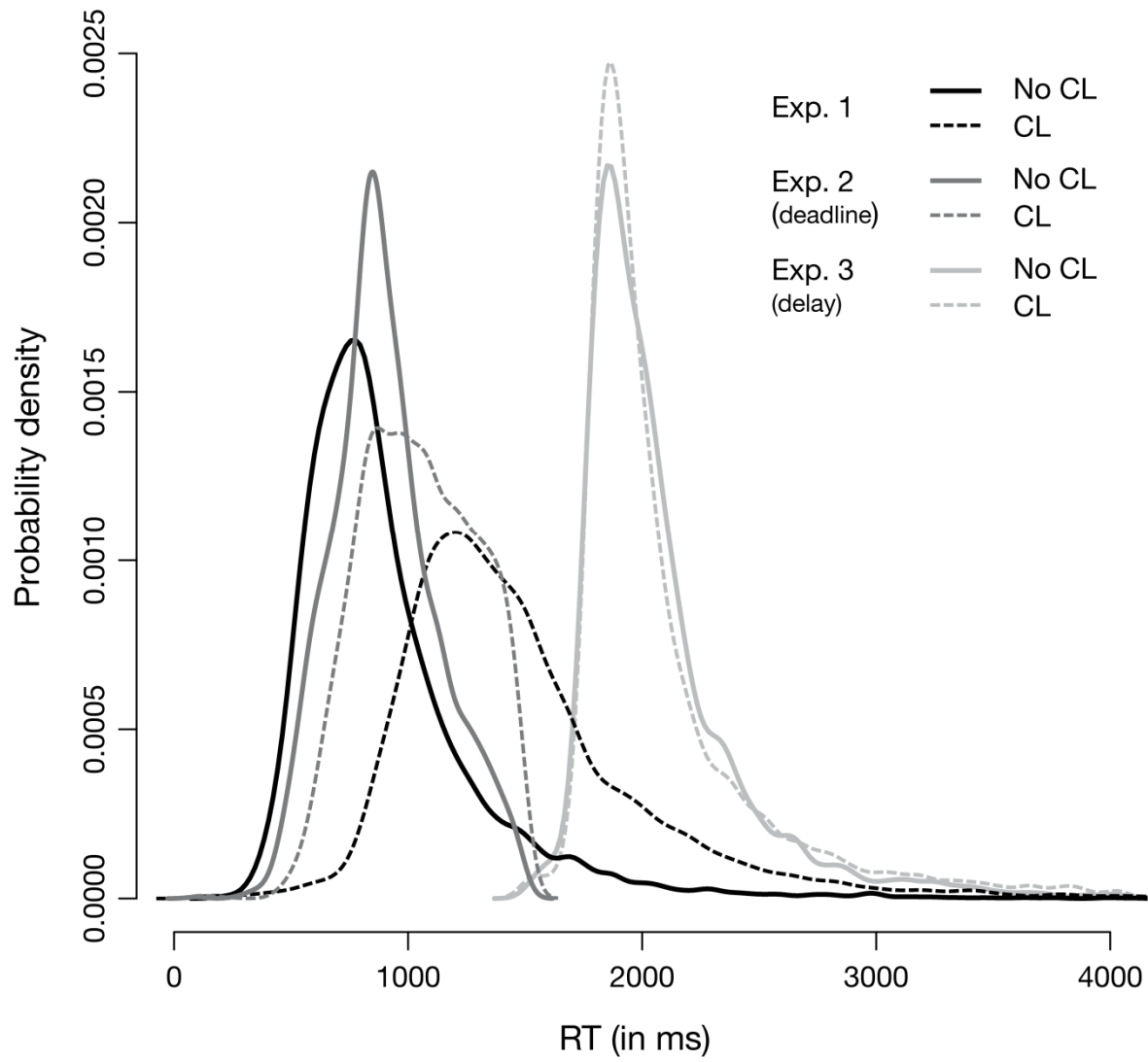


Figure 4.

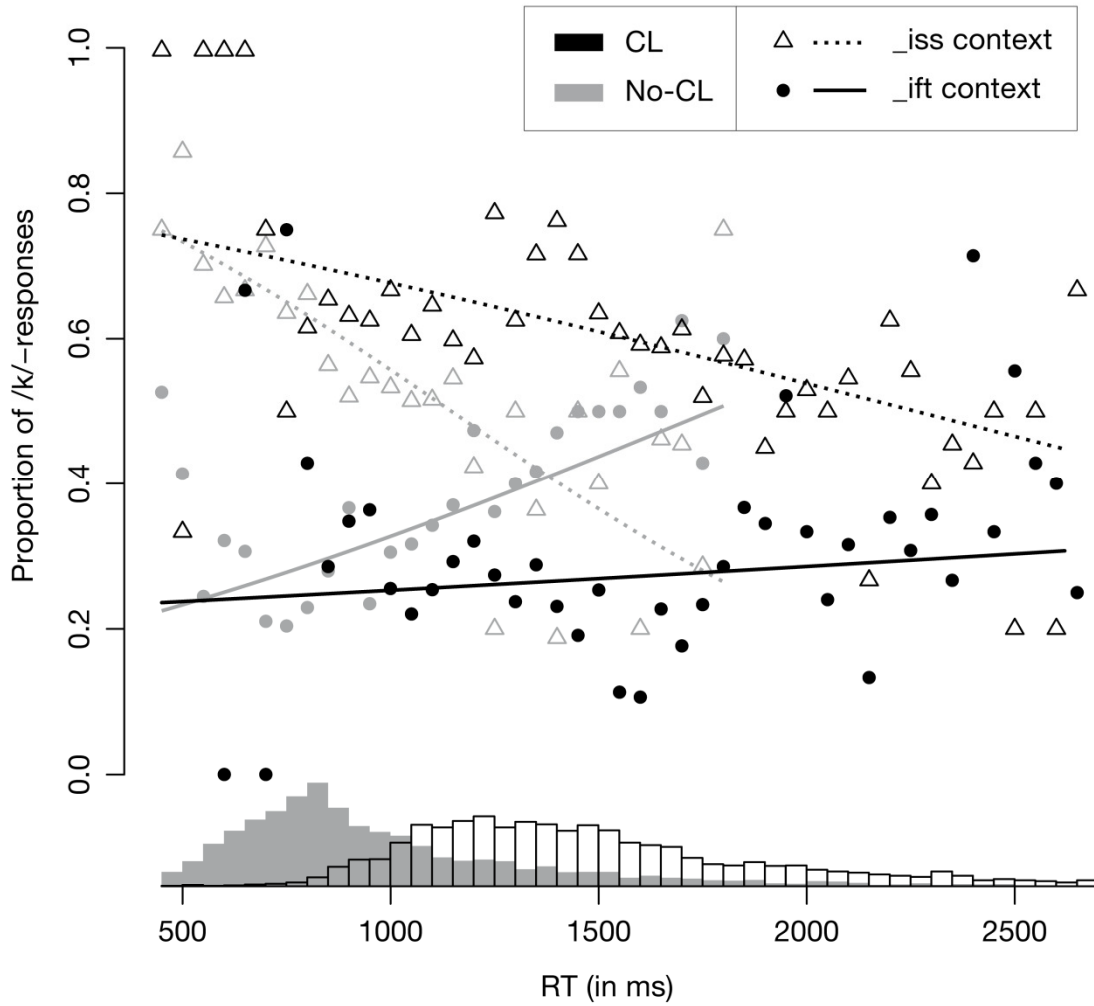


Figure 5.

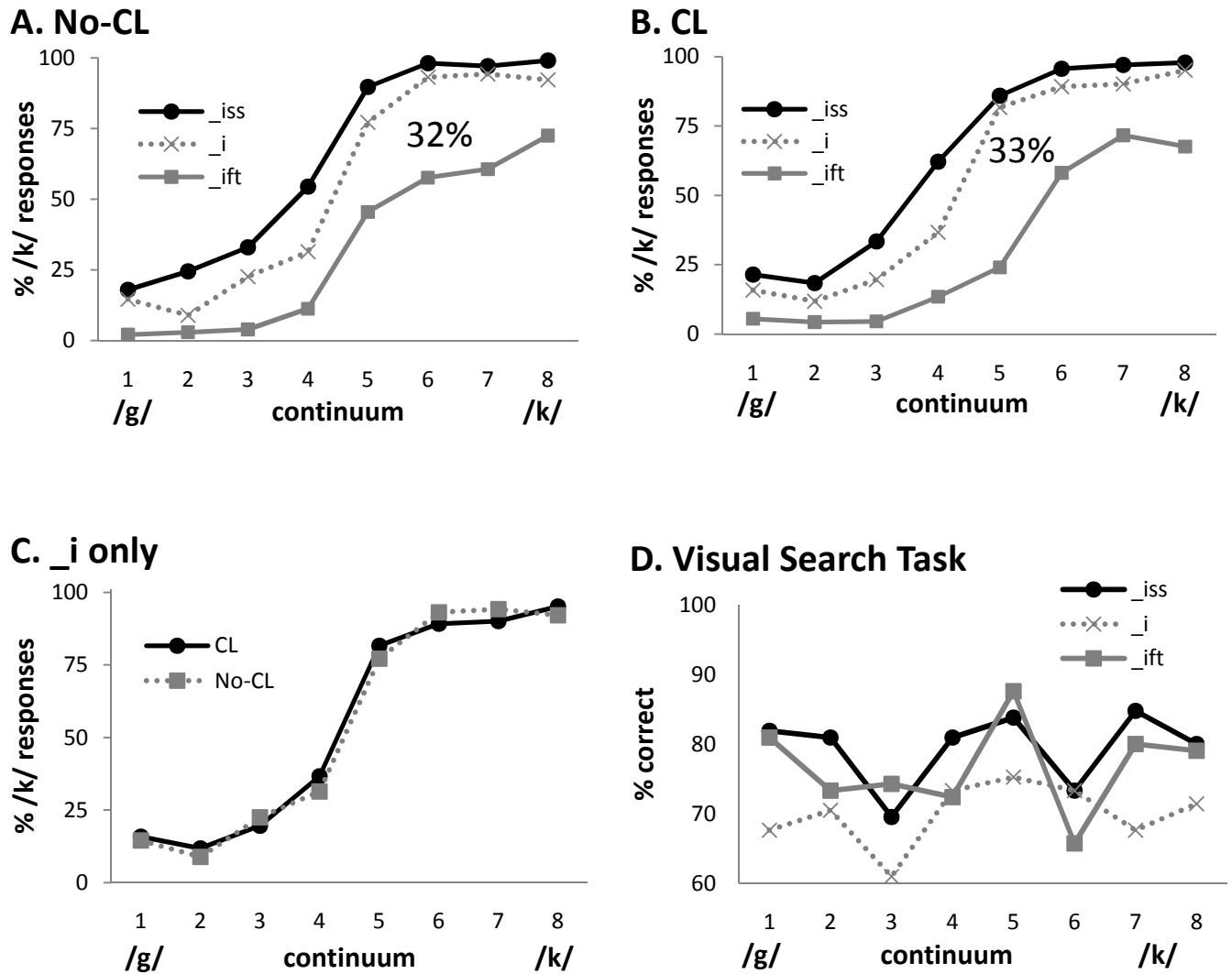
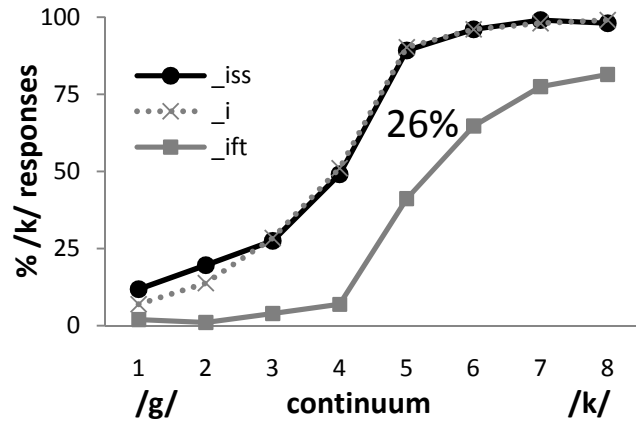
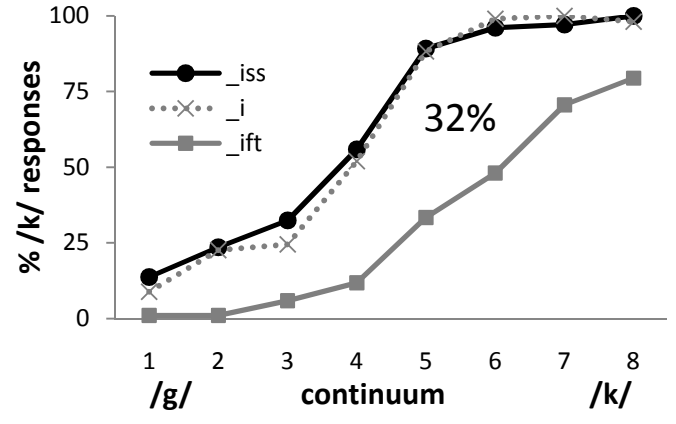


Figure 6.

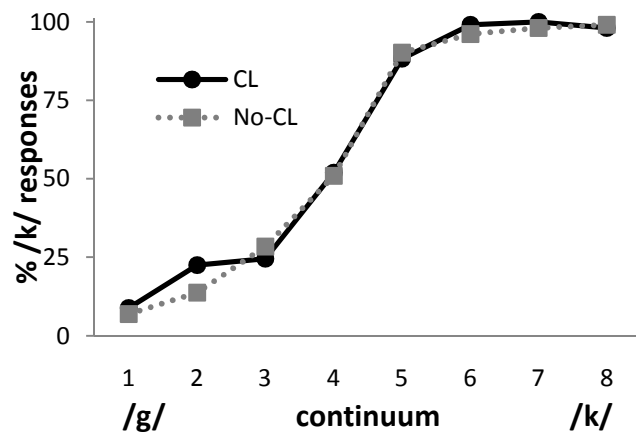
A. No-CL



B. CL



C. _i only



D. Visual Search Task

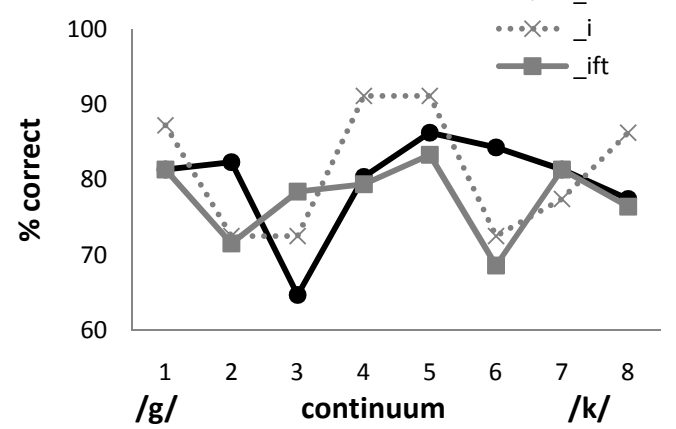


Figure 7.

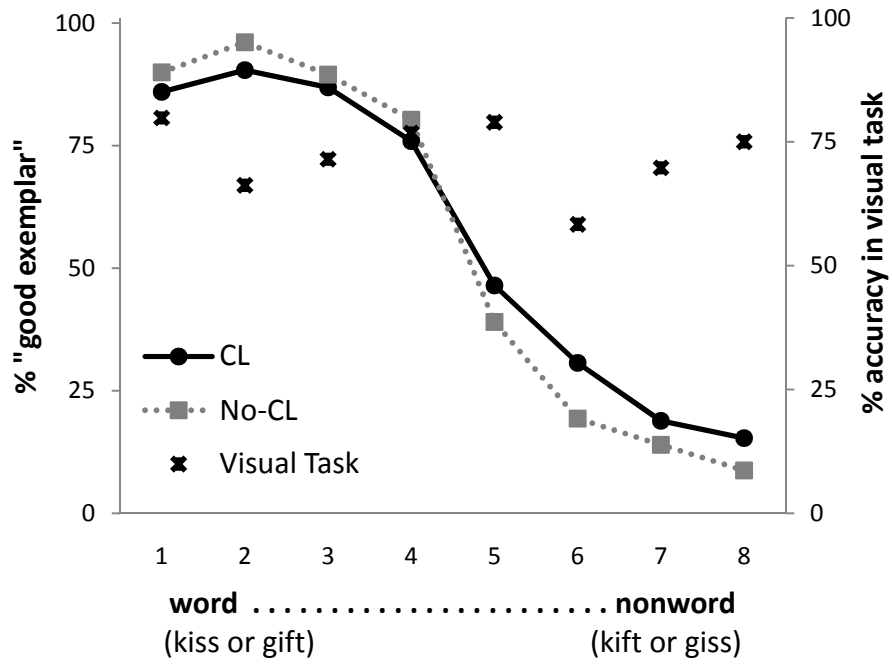


Figure 8.

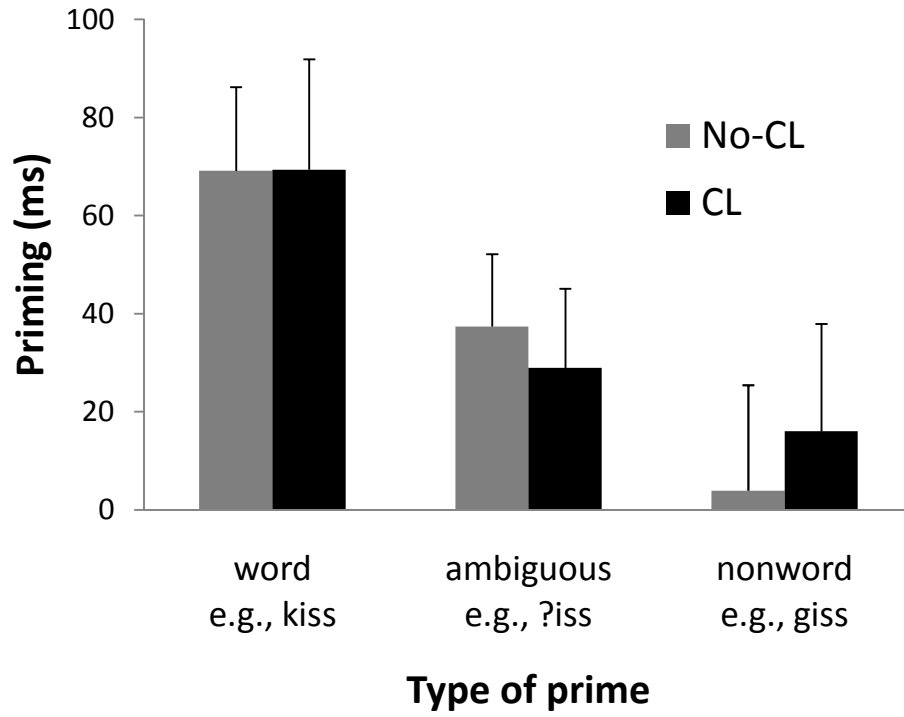
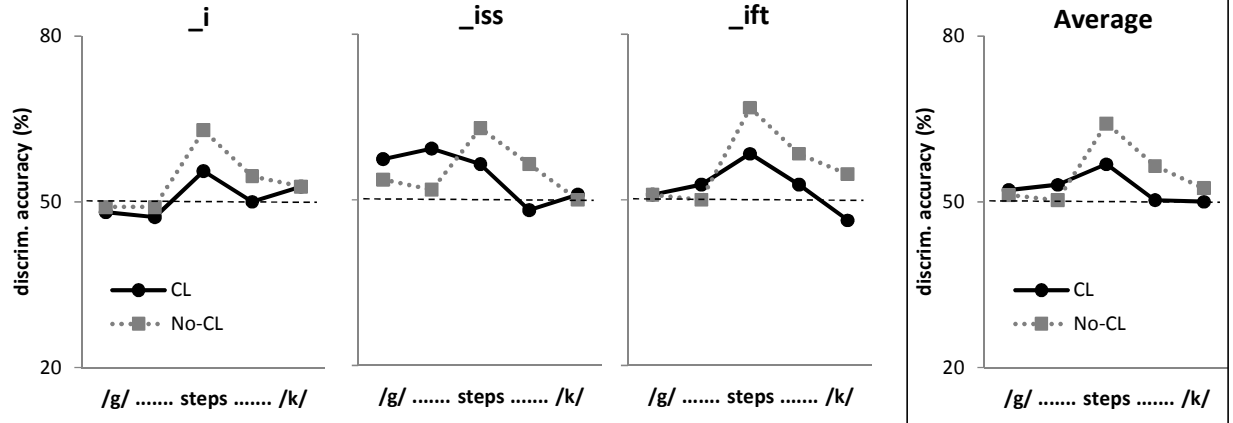


Figure 9.

One step



Two steps

