Published in final edited form as:

Psychol Res. 2008 November ; 72(6): 666–674. doi:10.1007/s00426-008-0166-2.

Visual Word Recognition is Accompanied by Covert Articulation: Evidence for a Speech-like Phonological Representation

Brianna M. Eiter and
Hofstra University

Albrecht W. Inhoff
State University of New York at Binghamton

Abstract

Two lexical decision task (LDT) experiments examined whether visual word recognition involves the use of a speech-like phonological code that may be generated via covert articulation. In Experiment 1, each visual item was presented with an irrelevant spoken word (ISW) that was either phonologically identical, similar, or dissimilar to it. An ISW delayed classification of a visual word when the two were phonologically similar, and it delayed the classification of a pseudoword when it was identical to the base word from which the pseudoword was derived. In Experiment 2, a LDT was performed with and without articulatory suppression, and pseudowords consisted of regular pseudowords and pseudohomophones. Articulatory suppression decreased sound-specific ISW effects for words and regular pseudowords but not for pseudohomophones. These findings indicate that the processing of an orthographically legal letter sequence generally involves the specification of more than one sound code one of which involves covert articulation.

Keywords

Reading; Phonological Representation; Covert Articulation; Visual Word Recognition; Eye Movements; Working Memory

All European writing systems use visual symbols, letters and letter clusters, to represent the phonemes of the corresponding spoken language. From its inception, reading research has sought to determine whether recovery of the represented sound code is part and parcel of visual word recognition. Subjectively, readers often experience hearing the words they read, and several recent word recognition experiments suggest that graphemes are converted into a speech-like code (Abramson & Goldinger, 1997; Birch, Pollatsek, & Kingston, 1998; Lukatela, Eaton, Sabadini, & Turvey, 2004). Tongue-twister sentences also are notoriously difficult to read even when reading is silent (McCutchen, Bell, France, & Perfetti, 1991; McCutchen & Perfetti, 1982), and the known speaking rate of ‘authors’ influences the time that participants spend reading their prose (Kosslyn & Matt, 1977).

Effects of speech properties on successful word recognition, such as effects of articulation duration on lexical decision times (Lukatela et al., 2004) have been attributed to the use of speech-like codes for visual recognition. Yet, overt speech, as occurs when reading out loud tends to slow reading and the eyes are typically one to several words ahead of an articulated word (e.g., Levin, Grossman, Kaplan, & Yang, 1972). Results from our laboratory (Inhoff,

Connine, Eiter, Radach, & Heller, 2004) suggest that a word's speech-like code is represented after it has been recognized in reading. In our study, participants read sentences with a phonologically regular target word while eye movements were recorded. An invisible spatial boundary was established immediately to the left of the target and the movement of the eyes that positioned the eyes to the right of the boundary initiated the presentation of an irrelevant spoken word (ISW) to the reader via headphones. The ISW for a visual target, e.g., "plate", was either identical to the target, "PLATE", phonologically similar to it, "PLACE", or dissimilar, "HORSE" (neither the phonologically similar companion word nor the dissimilar ISWs fit into sentence context). Examination of the consequences of the contingent speech presentation on visual target word and post-target word viewing revealed two distinct effect patterns, each of which was time-locked. During target reading, similar and dissimilar ISWs increased target viewing durations to the same extent relative to the identical ISW. That is, similar and dissimilar ISWs appeared to hamper target reading because their meaning disagreed with the meaning of the fixated visual target. During the reading of the following post-target word, by contrast, viewing durations were significantly longer in the similar ISW condition than the dissimilar condition. The recognition of a visual target word during reading was thus followed by a brief period during which the phonological –rather than the semantic-properties of an ISW context influenced reading. We attributed this phonological post-target reading effect to disruptive effects of the phonologically similar ISW on the phonological working memory representation of the previously identified visual target.

Yet, other recent studies also show rapid extraction of information for phonological code computation, within 30ms of item visibility (Lee, Binder, Kim, Pollatsek, & Rayner, 1999; Lee, Rayner, & Pollatsek, 1999; Perfetti & Bell, 1991). During reading, extraction of information for phonological code specification occurs at an earlier point in time than use of information for the specification of word meaning (Lee et al, 1999), and phonological properties of a target word can be extracted before it is viewed (Pollatsek, Lesch, Morris, & Rayner, 1992; Liu, Inhoff, Ye, & Wu, 2003; see also Pollatsek, Tan, & Rayner, 2000). The relatively early point in time at which phonology-defining information is extracted during visual word recognition has been used to argue that these codes are impoverished, potentially consisting of sub-word segments, that they are relatively abstract, and that they are devoid of speech-type features (Frost, 1998).

Rather than contradicting each other, these different time lines can be reconciled when one tacitly held assumption is abandoned, i.e., that readers compute just one effective sound code. Without this assumption, computation of one sound code could occur rapidly and yield an impoverished code that is devoid of speech-like features, and a somewhat slower computation could yield a more detailed sound code that represents speech-like features. This speech-like code may persist after word recognition in phonological working memory. Alternatively, readers could use a single, relatively abstract phonological code for visual word recognition and another speech-like code for phonological working memory representation. In Baddeley's influential model of working memory, recognized visual words must undergo a process of covert articulation before they can be represented in phonological working memory (Baddeley & Hitch, 1972).

The current study examined two related issues. Experiment 1 sought to determine whether sound-specific effects of an ISW that emerged immediately after target reading in an earlier study (Inhoff et al., 2004) interfered with the target's recognition or with its representation in phonological working memory. A lexical decision task, rather than a sentence reading task, was used for that purpose. The task can be performed without the explicit overt use of phonological codes and, critically, it does not involve the representation of an item in working memory. A lexical decision response is performed in response to a visually depicted item and the item is no longer task relevant once the classification has been performed. In silent reading,

by contrast, word recognition must be followed by the representation of identified words in working memory, so that higher-order phrase and sentence meaning can be determined. If the phonologically similar ISW interfered with a visual target word's recognition, then sound specific ISW interference should be present in the lexical decision task. Conversely, if the phonologically similar ISW interfered with the target's working memory representation during reading, then performance in the LDT should not be subject to sound-specific ISW effects. Experiment 2 sought to determine the nature of the phonological code that is used for lexical discrimination, i.e., whether it is relatively abstract and devoid of speech-like features or whether it resembles covert speech.

Experiment 1

Method

Participants—Forty-eight undergraduate students from Binghamton University participated in the experiment in order to fulfill a requirement for a Psychology course. All had normal or corrected to normal vision and no participant reported any difficulty hearing the auditory presentations.

Materials—The experimental stimuli consisted of fifty-six items; forty-two target words and fourteen filler words. All words were either five or six letters long, and visual target words were nouns with a mean word frequency of 17 per million with a range from .5 to 129 per million (according to the CELEX database, Baayen, Piepenbrock, & Van Rijn, 1993). Thirty target words were monosyllabic; the remaining 12 were bisyllabic. The identical set of words was used in a prior reading study where these target words were read within neutral sentence contexts (Inhoff et al., 2004).

Forty-two pseudowords were created from source words, all nouns, that ranged in length from 5 to 7 letters and had a mean word frequency of 2.65 per million, with a range from .11 to 4.8 per million. Pseudowords were created by replacing one letter, in either the second, third, fourth, or fifth position, with a letter that made the resulting letter string an orthographically regular, pronounceable, but nonmeaningful letter sequence. For instance, to create a pseudoword the letter I in the word HINGE was replaced with an O to create the pseudoword HONGE. As in Inhoff et al. (2004), three different types of ISW were assigned to each visual target word, one that was identical in phonological form and meaning to the word, one that was similar in phonological form and dissimilar in meaning, and one that was dissimilar in phonological form and meaning. For instance, the visual target 'plate' was presented with the irrelevant spoken word PLATE in the identical condition, with PLACE in the similar condition, and with HORSE in the dissimilar condition. Three different types of ISW were also paired with each pseudoword, one that was identical in phonological form and meaning to the source word from which these items was derived, the base-identical condition, one that was similar in phonological form and dissimilar in meaning, the similar condition, and one that was dissimilar in both phonological form and meaning. For instance, the pseudoword 'honge' was paired with the ISW HINGE in the base-identical condition, with HITCH in the phonologically similar condition, and with FLUTE in the dissimilar condition. (Base-)Identical and dissimilar ISWs were matched for syllable number but did not share any phonemes or letters at corresponding word locations. To create a dissimilar ISW condition, identical spoken words were paired with visual targets that did not have phonological or semantic overlap.

A native speaker of American English, trained in the production of canonic word forms, articulated the ISWs at a normal tempo for word and pseudoword items. A directional microphone (Optimus 33–3017) and speech editing software (Soundforge 4.5) were used to record words at 22,050 Hz in 16-bit stereo. The duration of ISWs ranged from approximately 350 ms to 700 ms, with a mean of 546 ms for identical ISWs and of 534 ms for similar ISWs.

The duration of the ISWs paired with pseudowords ranged from approximately 350 ms to 900 ms, with a mean of 625 ms for the identical condition, and 602 ms for the similar and dissimilar condition. Spoken words were presented over a pair of high quality (Beyerdynamic DT 250) headphones. All acoustic signals were presented at a comfortable loudness level of approximately 70 db.

Apparatus—Stimuli were presented on a CRT in white on a black background and lexical decision responses were recorded via a response box interfaced with a PC which was used to control the timing of the visual display and to record lexical decision latencies (i.e., the interval between the onset of the target word and the activation of the key press). All lexical decision latencies were measured to the nearest millisecond. Head and chin position were not restrained, and the CRT was positioned to yield an eye-stimulus distance of approximately 40 cm. At this distance, each character subtended approximately .45 degrees of visual angle horizontally and .5 degrees vertically.

Procedure—All participants were tested individually. Each trial began with the presentation of a fixation marker at the screen center for 500 ms. Participants were instructed to fixate the marker and they were informed that a word or a pseudoword would appear at the fixated location. The fixation marker was followed by a 100 ms blank interval, the subsequent presentation of the visual target (either a word or pseudoword), and the onset of the ISW. The visual stimulus remained on the screen for 800 ms. Participants were instructed to decide whether the target was or was not a word of the English language by pressing on the right button if it was a word and the left button if it was not a word. Ten practice trials preceded the experiment.

Design—Three lists were used, all of which contained the identical sets of 42 target words from Inhoff et al. (2004). In addition each list contained the identical set of 42 regular pseudowords. Each visual item was paired with a different ISW across lists, and 14 words and 14 pseudowords were presented in each of the three ISW conditions on each list. The ordering of ISW contexts across consecutive trials and of word and pseudoword trials was randomized. The assignment of lists was counterbalanced across three consecutive participants.

Responses made less than 250 ms or more than 1500 ms after the onset of the visual target were excluded from the analysis of lexical decision latencies. This removed approximately 2.6% of the raw data. Errors, consisting of incorrect responses, were also excluded from the overall analysis. This removed another 10% of the raw data from the RT analyses. Analyses of variance (ANOVA) with the within subjects factor ISW context (identical, similar, dissimilar) were applied to RTs and error rates and both subject and items analyses (F_1 and F_2) are reported. Subanalyses that contrasted the similar condition with the identical and dissimilar conditions were applied when the main effect of ISW context was reliable.

Results

Mean correct lexical decision times (LDTs) and percentage of errors for words and pseudowords are shown in Table 1. As can be seen, ISW type influenced LDTs for target words, $F(2,94) = 9.96$, $MSE=2612$, $p<.01$; $F(2,82) = 7.65$, $p<.01$, $MSE=4376$. Paired contrasts that used the dissimilar condition as baseline revealed a significant 24 ms benefit for identical ISWs, $t_1(47) = 2.90$, $p<.025$; $t_2(41) = 4.91$, $p<.03$, and a marginally reliable 22 ms disadvantage for similar ISWs, $t_1(47) = 1.93$, $p<.06$; $t_2(47) = 2.40$, $p = .1$.

The main effect of ISW type was not reliable in the analysis of pseudowords, $F_1(2,94) = 1.99$, $MSE=1953$, $p<.15$; $F_2(2,82) = .22$, $p = .80$, $MSE=9311$. Nevertheless, an inspection of mean values indicates an effect pattern that is consistent with interference caused by phonological

similarity. RTs were *longest* in the base-identical condition, where the ISW matched the phonological form of the base word, i.e., when the phonological similarity between the ISW and the corresponding pseudoword was relatively high, RTs were intermediate in the similar condition, and shortest in the dissimilar condition. The 18 ms difference between the base-identical and the dissimilar condition was reliable in the subjects analysis, $t_1(47) = 2.38$, $p < .025$, but not in the items analysis, $t_2(41) = .60$.

ISW also influenced the error rates for words, $F_1(2,94) = 23.22$, $MSE = .004$, $p < .01$; $F_2(2,82) = 9.74$, $p < .01$, $MSE = 1.94$. Using the dissimilar condition with a 10% error rate as baseline, errors were decreased by 3% in the identical condition, $t_1(47) = 2.38$, $p < .025$; $t_2(41) = .18$, $p = .86$. and increased by 7% in the similar condition, $t_1(47) = 4.60$, $p < .01$; $t_2(41) = 1.58$, $p = .1$. ISW did not influence pseudoword classification, F_1 and $F_2 < 1$. Overall, pseudowords error rate was 9% and the three ISW conditions differed by less than 2%.

Discussion

The lexical classification of words was more difficult when their visual presentation was accompanied by the presentation of a phonologically similar ISW than when presentation was accompanied by an irrelevant phonologically identical or an irrelevant dissimilar ISW. The corresponding negative similarity effect was marginally reliable in the LDT data and highly reliable in error data. The LDT data are thus in agreement with the results of an earlier reading study (Inhoff et al. 2004) in which the same target words and ISWs generated sound-specific interference immediately after the visual target was viewed. LDTs for pseudowords also revealed a tendency toward a negative effect of phonological similarity in this task, as classification was relatively slow in the base-identical condition where the ISW and the to-be-classified pseudoword were highly similar.

The present results suggest a different account of the negative phonological similarity effect as it occurred in Inhoff et al. (2004). Rather than attributing negative effects of phonologically similar irrelevant speech to disruptions of phonological working memory representations, the current data suggest these effects occur before a working memory representation is established. In contrast to reading where identified items must remain active in working memory so that sentence meaning can be determined, working memory should not be an effective contributor to the success of an ongoing lexical decision task because there is no need to maintain a phonological representation.

While it is plausible that the disruptive effects of a phonologically similar ISE in the lexical decision task indicate that phonological similarity of irrelevant speech impeded the success with which a visual item accessed a lexical representation an alternative account of these findings cannot be ruled out. Specifically, the negative effect of a phonologically similar ISW could occur immediately after the word has been accessed, e.g., during a post-access verification process (Balota & Chumby, 1985). Previous research suggests that the lexical decision task may involve more than one stage of processing: an identification process, and also a post-access verification process (Balota & Chumby, 1985). In the lexical decision task, post-access verification influences reaction times when processing demands are great. In the current experiment, making lexical decisions while concurrent irrelevant speech is presented increases the processing demands on the participant, thus making it likely the post-access verification process was involved in decision making. Consequently, the negative phonological similarity effect is likely to have occurred during the later process and not during prelexical and/or lexical processing. Therefore, the results of Experiment 1 suggest that the phonologically similar ISW was interfering with a phonological or articulatory process that survives lexical access and precedes working memory storage.

Related work indicates that word recognition can be accompanied by covert articulation. In a recent study, Saito and Baddely (2004) showed that irrelevant sound could induce phonologically related slips of the tongue. Brain imaging techniques further showed that mere listening to speech facilitated tongue and lip muscles (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Watkins, Strafeela, & Paus, 2003), and that it activated motor areas that are involved in speech production (Wilson, Saygin, Sereno, & Iacoboni, 2004). According to these findings, listening to an ISW in Experiment 1 may have automatically activated corresponding speech-motor areas. Lexical classification of visual items could also have activated speech areas if processing was accompanied by their covert articulation. Sound-specific interference for words may have occurred because the motor area activated by the phonologically similar ISW overlapped with the motor area that was also used for the covert articulation of a different word, the visual target. Similarly, sound specific interference occurred for pseudowords in the base-identical condition because the motor area activated by the legal ISW overlapped with the motor area activated by the pseudoword. 'No' responses to pseudowords could have been particularly difficult under these circumstances.

The classification of words also revealed a robust identity advantage, with shorter LDTs and fewer errors in the identical than the dissimilar ISW condition. A similar identity advantage was present during target viewing in the earlier reading Experiment (Inhoff et al., 2004). Experimental effects during reading across -consecutive words, the target and the post-target word, thus emerged during the classification of a single word, the target, in the lexical decision task. LDTs are generally much longer than comparable viewing times for words during reading, and the relatively long latency in the LDT could provide an opportunity for the influence of relatively early and of relatively late processes on performance in this task.

To examine the time line of identity and phonological ISW effects, a median split was applied to LDTs in the identical condition (for words) to create two equal-size groups of participants, one with relatively short LDTs (597 ms) and one with relatively long LDTs (821 ms). Using the dissimilar condition as baseline, the fast groups yielded a larger identity advantage, 35 ms, than the slow group, 17 ms. Conversely, the fast group yielded smaller negative effects of ISW similarity, 10 ms, than the slow group, 35 ms. These results are consistent with the view that the identity effect preceded the negative sound similarity effect. Together, the results of Experiment 1 thus suggest that visual items were covertly articulated after lexical access was either achieved or attempted.

Experiment 2

Experiment 2 sought to test the assumption that the lexical discrimination of a visual item was accompanied by its covert articulation, which we refer to as the covert articulation hypothesis. Toward this goal, the LDT task of Experiment 1 was performed under two conditions, one in which the task was performed exactly as in Experiment 1, and one in which it was performed while articulatory suppression was in progress. Articulatory suppression hinders covert articulation (Baddeley, 1986), and it should consequently diminish sound-specific ISW effects if they are due to the covert articulation of the visual item. If item recognition involved covert articulation, then the negative phonological ISW effect for words and the negative base-identity effect for pseudowords should be diminished when articulatory suppression is added to the experimental task.

Experiment 2 also included a second test of the 'covert articulation' hypothesis. Specifically, we manipulated the type of pseudoword so that half were regular pseudowords and half were pseudohomophones. The phonological form of a pseudohomophone hampers lexical discrimination, and in the absence of articulatory suppression lexical discrimination of these items is generally slower and more error prone than classification of regular pseudowords

(Berent, 1997; Ferrand & Grainger, 1996; Pexman, Lupker, & Jared, 2001; Stone & Van Orden, 1993), which is often referred to as the pseudohomophone effect. This effect should be smaller with articulatory suppression if suppression hinders the specification of an effective sound code.

Method

Participants—Forty-eight undergraduate students from Binghamton University who were not familiar with the experimental materials participated in the experiment in order to fulfill a course requirement. All had normal or corrected to normal vision and none reported any difficulty identifying the auditory ISW presentation.

Materials, Apparatus, and Procedure—The same apparatus was used as in Experiment 1, and the target words of Experiment 2 were identical to the target words of Experiment 1. The pseudoword materials differed, however, across experiments. Specifically, half (24) of the pseudowords in Experiment 2 were pseudohomophones and half were regular pseudowords that were taken from Experiment 1. As was the case for regular pseudowords, the graphemic form of the pseudohomophone differed from the graphemic form of the base word by one letter, e.g., the pseudohomophone ‘sleuth’ was derived from the base word ‘sleuth’. Again, regular pseudowords and pseudohomophones were paired with the phonological form of the base word in the base-identical condition, SLEUTH, with a word that was similar to the phonological form of the base word in the identical condition, SLEIGH, and with a different word in the dissimilar condition, THRILL.

The no articulatory suppression condition of Experiment 2 was identical to the experimental task of Experiment 1, in that it required the discrimination of words from pseudowords. In the articulatory suppression condition, participants were asked to repeat the spoken letter sequence A, B, C, D, E, and F out loud throughout the experiment. They were instructed to repeat the letter string at their own pace, and to continue repeating the letter string throughout and in between experimental trials. Two practice sessions with 10 items each preceded the experiment.

Three lists were created with the same target words and pseudowords. Similar to Experiment 1, each list was paired with a different ISW and list assignment was counterbalanced across three consecutive participants. The order of ISW conditions was randomized, the order of regular pseudowords and pseudohomophones was not. Instead, 80% of all pseudowords were presented during the first half of the experiment to ensure that all participants were aware that pseudohomophones were included on the list so that readers would focus on the use of orthographic information in the classification task.

The same criteria was used to eliminate outliers (3.4%) and errors from RT analyses in Experiment 2 as in Experiment 1. ANOVAs were applied to latencies and error rates with the between subjects factor articulatory suppression (present yes absent), and the two within subjects factors ISW context (identical, phonologically similar, and dissimilar) and pseudoword type (regular pseudoword vs pseudohomophone).

Results

LDTs and error rates for target words and for the two types of pseudowords are shown in Table 2.

LDTs for words tended to be shorter with than without articulatory suppression, although this trend in the data was not statistically significant for subjects analyses, $F(1,46) = 2.43$, $MSE=64314$, $p<.13$, it was for items analyses $F(1,82) = 9.95$, $p<.01$, $MSE=53720$. The main

effect of ISW was reliable, $F(2,92) = 21.626$, $MSE=1249$, $p<.01$; $F(2,164) = 30.82$, $p<.01$, $MSE= 1547$ again with relatively short RTs in the identical condition and relatively long RTs in the similar condition. The main ISW effect was influenced by articulatory suppression, however, and the effect pattern corresponded to that predicted by the covert articulation hypothesis. Specifically, without articulatory suppression, LDTs were 29 ms longer in the similar than in the dissimilar condition, an effect size that is slightly larger than in Experiment 1. With articulatory suppression, by contrast, LDTs were 5 ms shorter in the similar than the dissimilar ISW condition. The corresponding interaction of the similar and dissimilar ISW with articulatory suppression was reliable, $F(1,46) = 5.37$, $MSE=1249$, $p<.05$; $F(1,82) = 7.41$, $p<.01$, $MSE = 1643$. The identity advantage showed an opposite influence of articulatory suppression with a larger advantage with than without suppression, 21 ms and 47 ms, respectively; the corresponding interaction was marginally reliable in the subjects analysis $F(1,46) = 4.04$, $MSE=1070$, $p<.051$, but not significant in the items analysis $F(1,82) = .33$, $p = .56$, $MSE=2574$.

Pseudoword decisions were also shorter, by 40 ms, with than without articulatory suppression. Although this effect was sizable, it did not approach significance in the subjects analyses, $F(1,46) < 1$, due to large RT variability in the data, however it did in the items analyses, $F(1,82) = 19.26$, $p<.01$, $MSE= 80692$. As predicted, both types of pseudowords showed a smaller identity disadvantage (the difference between the base-identical and dissimilar pseudowords conditions) with than without articulatory suppression, 62 ms and 38 ms, respectively, but the statistical two-way interaction of ISW type and articulatory suppression did not approach significance, $F=1$. Overall, categorization of pseudohomophones was relatively time consuming in the base-identical condition and the corresponding two-way interaction was reliable, $F(2,92) = 6.15$, $MSE=5089$, $p<.01$; $F(2,164) = 19.26$, $p<.01$, $MSE= 80692$.

LDTs for pseudowords were also examined. They yielded a robust pseudohomophone disadvantage, $F(1,46) = 4.23$, $MSE=8759$, $p<.05$; $F(1,82) = 4.23$, $p<.04$, $MSE=48206$. As predicted, the pseudohomophone effect was smaller with articulatory suppression, 8 ms, than without suppression 37 ms; this sizable numeric effect was not statistically reliable in the F_1 analysis, $F(1,46) = 1.75$, $MSE=15335$, $p<.2$; but it was robust in the F_2 analyses, $F(1,82) = 5.41$, $p<.025$, $MSE=101771$.

Mean errors for words and pseudowords are also shown in Table 2. Error rate for words was approximately 10% across conditions, and there were no reliable differences, all F_1 and $F_2 < 1$. Pseudoword classification was more error prone with than without articulatory suppression, 17% and 12%, respectively, although the corresponding main effect did not reach significance, $F(1,46) = 2.65$, $MSE=5.4$, $p<.12$; $F(1,82) = 1.33$, $p<.25$, $MSE=2.76$. Classification of pseudohomophones was more error prone, 20%, than classification of regular pseudowords, 10%, $F(1,46) = 39.59$, $MSE=5.4$, $p<.01$, $F(1,82) = 39.59$, $p<.01$, $MSE=5.4$ and the difference in error rate between the two pseudoword types was particularly large in the base-identical condition. The corresponding two-way interaction of pseudoword type and ISW was reliable, $F(2,92) = 4.0$, $MSE=1.5$, $p<.025$; $F(2,164) = 8.75$, $p<.01$, $MSE= .85$.

Discussion

Articulatory suppression had a profound influence on the ISW effect for words. A negative effect of phonological ISW similarity in the standard lexical decision task ceased to exist when articulatory suppression was added to the experimental task. Articulatory suppression was expected to hinder the covert articulation of the visual target word. The phonologically similar ISW thus no longer hampered the classification of the visual target word because the two different words no longer competed for overlapping speech-motor areas. Without an effective phonological representation of the visual target, that is established by covert articulation, task-relevant orthographic and lexical properties of the visual item may become more prominent

thus yielding overall somewhat faster responses, and particularly fast and accurate responses for words in the identical condition where lexical representation of the visual target matched the lexical representation of the ISW.

The pseudoword data yielded a more heterogeneous pattern of effects. In contrast to words which were identical to the words used in Experiment 1, the pseudowords of Experiment 1 were split into two classes, regular pseudowords and pseudohomophones, thus splitting the number of items per condition in half. Overall, the pseudoword data were thus subject to much more variability so that even sizable trends in the data did not reach statistical significance. Two such trends in the pseudoword data are in harmony with the predictions of our covert articulation hypothesis. The negative base-identity effect was diminished with articulatory suppression, and ISW effects in general, disappeared for regular pseudowords in the articulatory suppression condition, as should be the case if articulatory suppression hindered the specification of a phonological code for visual targets. Moreover, the magnitude of the pseudohomophone effect was greatly diminished with articulatory suppression.

However, pseudohomophones also yielded two findings that appear to pose difficulties for the covert articulation hypothesis. Irrespective of articulatory suppression, RTs were relatively long in the base-identical condition. Furthermore, errors were also more common for pseudohomophones than for regular pseudowords and this pseudohomophone effect was of approximately the same magnitude with and without articulatory suppression. The implications of these two findings are considered in the General Discussion. (I suggest that the next part be deleted because the list composition was the same when the task was performed with and without articulatory suppression).

General Discussion

The current data revealed sizable and statistically robust effects of an ISW on the lexical classification of words and pseudowords. In Experiment 1, LDTs for the identical set of words was slower when they were presented in the context of a phonologically similar ISW than when they were presented in the context of a dissimilar ISW. The LDT data also revealed that ISWs that were identical to the visual target word facilitated classification relative to dissimilar ISWs. These data are consistent with results from an earlier reading study (Inhoff et al., 2004), where an ISW yielded a positive identity effect and a negative phonological similarity effect during the reading of two consecutive words: an identical ISW facilitated target reading, relative to the baseline, and a similar ISW hampered the reading of the next word(s) in the text. They extend our earlier findings in that they show that deleterious effects of a phonologically similar ISW need not arise from interference between the phonological forms of the ISW and an identified visual word in phonological working memory. Instead, a phonologically similar ISW appears to interfere with a process that follows the accessing –or attempted accessing– of a lexical representation irrespective of working memory demands.

To examine the nature of a visual word's post-recognition process, the LDT was performed with and without concurrent articulatory suppression in Experiment 2. The negative phonological similarity effect for words disappeared with articulatory suppression, as predicted by the covert articulation hypothesis, according to which processing of a visual word or pseudoword is accompanied by its covert articulation. Since mere listening to an ISW activates corresponding motor speech representation, the covert articulation of the phonologically similar ISW and the target word could impede processing because this ISW and the visual target used highly overlapping speech-motor areas that were connected to **different** lexical representations. Consequently, the selection of one particular lexical representation, that matched the visual target but not the ISW, was likely to be slow and error prone under these conditions.

Articulatory suppression did not only involve overt speech activity it also generated an acoustic background within which the LDT was to be performed. It could thus be argued that our key evidence for the covert articulation hypothesis, the cessation of the negative phonological similarity effect with articulatory suppression, suffers from a confounding of articulatory and acoustic effects. That is, the acoustic background generated by articulatory suppression may have leveled differences between ISW types, thus eliminating the negative phonological similarity effect. One key aspect of the data argues, however, against this account. The effect of articulatory suppression was highly selective. It did not eliminate the advantage of the identical ISW condition over the dissimilar ISW condition; to the contrary, this advantage was numerically larger with articulatory suppression.

Although the results of Experiments 1 and 2 converge in showing that visual target items were covertly articulated after lexical access was either achieved or attempted, the results cannot be taken as evidence for the view that covert articulation accounts for all sound-specific coding of visual target items. Two findings from Experiment 2 argue against such a generalized account. Specifically, articulatory suppression did not diminish deleterious effects of the ISW in the base identical condition irrespective of articulatory suppression, and it did not decrease the higher error rate for pseudohomophones than for regular pseudowords. If articulatory suppression completely inhibited the covert articulation and the phonological representation of to-be-classified pseudohomophones then their sound-specific properties should have become ineffectual. The two pseudohomophone effects thus converge, indicating that classification of these items must have involved use of phonological knowledge that was not subject to the effect of articulatory suppression.

Covert articulation of a visual target appears to occur relatively late during target processing and covert articulation interferes with that. The recognition of a visual target may also involve the computation of a relatively abstract phonological code that is not subject to articulatory suppression, and computation of this code may precede the accessing of lexical representations. Pseudohomophones differ from regular pseudowords in that their articulatory and abstract phonological codes match that of a regular word, rendering the word nonwords distinction more difficult and error prone, irrespective of articulatory suppression. Furthermore, articulatory suppression may not completely eliminate all covert articulation. If this was the case, then even a relatively marginal covert articulation of the pseudohomophone could increase the difficulty of the lexical decision task as it provides further evidence for a lexical status of the to be classified visual pseudohomophone.

Together, the results of Experiments 1 and 2 thus provide converging evidence for the obligatory covert articulation of a visual target item immediately after lexical access has been accomplished or attempted in the lexical decision task. Furthermore, presence of pseudohomophone effects irrespective of articulatory suppression in Experiment 2 provide preliminary evidence for the use of another functionally distinct phonological code that may be computed relatively early during visual word recognition. In her extensive review of the phonological working memory literature, Wilson (2001) discerned two anatomically distinct types of phonological representations, one that involves representation of speech-like features and one that does not. Our conception of the results of Experiments 1 and 2 is consistent with that.

Acknowledgements

This project was supported by NIH grant HD043405. We would like to thank Cynthia Connine, Marianne Lloyd, and two anonymous reviewers for their helpful comments on an earlier version of the paper. Inquiries should be directed to Brianna M. Eiter, Department of Psychology, State University of New York at Binghamton, Binghamton, N.Y. 13902-6000, USA.

References

- Abramson M, Goldinger SD. What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics* 1997;59:1059–1068. [PubMed: 9360478]
- Baddeley, AD. New York: Oxford University Press; 1986. *Working Memory*.
- Baddeley AD. Working memory: The interface between memory and cognition. *Journal of Cognitive Neuroscience* 1992;4:281–288.
- Baddeley, AD.; Hitch, GJ. Working Memory. In: Bower, G., editor. *The psychology of learning and motivation*. Vol. 8. New York: Academic Press; 1974. p. 47-90.
- Berent I. Phonological priming in the lexical decision task: Regularity effects are not necessary evidence for assembly. *Journal of Experimental Psychology: Human Perception and Performance* 1997;23:1727–1742. [PubMed: 9425677]
- Besner D. Phonology, lexical access in reading, and articulatory suppression: A critical review. *Quarterly Journal of Experimental Psychology* 1987;39A:467–478.
- Besner D, Davelaar E. Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology* 1982;36:701–711.
- Birch S, Pollatsek A, Kingston J. The nature of the sound codes accessed by visual language. *Journal of Memory and Language* 1998;38:70–93.
- Fadiga L, Craighero L, Buccino G, Rizzolatti G. Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience* 2002;15:399–402. [PubMed: 11849307]
- Ferrand L, Grainger J. List context effects on masked phonological priming in the lexical decision task. *Psychonomic Bulletin and Review* 1996;3:515–519.
- Frost R. Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin* 1998;123:71–99. [PubMed: 9461854]
- Inhoff AW, Connine C, Eiter BM, Radach R, Heller D. The phonological representation of words after their fixation during sentence reading. *Psychonomic Bulletin & Review* 2004;11:320–325. [PubMed: 15260200]
- Kosslyn SM, Matt AM. If you speak slowly, do people read your prose slowly? Person-particular speech recoding during reading. *Bulletin of the Psychonomic Society* 1977;9:250–252.
- Lee H, Rayner K, Pollatsek A. The time course of phonological, semantic, and orthographic coding in reading: Evidence from the fast-priming technique. *Psychonomic Bulletin & Review* 1999;6:624–634. [PubMed: 10682205]
- Lee Y, Binder KS, Kim J, Pollatsek A, Rayner K. Activation of phonological codes during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance* 1999;25:948–964. [PubMed: 10464940]
- Levin H, Grossman J, Kaplan E, Yang R. Constraints and the eye-voice span in right and left embedded sentences. *Language and Speech* 1972;15:30–39. [PubMed: 5073933]
- Liu W, Inhoff I, Ye Y, Wu C. Use of parafoveally visible characters during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance* 2003;28:1213–1227. [PubMed: 12421066]
- Lukatela G, Eaton T, Sabadini L, Turvey MT. Vowel Duration Affects Visual Word Identification: Evidence That the Mediating Phonology Is Phonetically Informed. *Journal of Experimental Psychology: Human Perception and Performance* 2001;30:151–162. [PubMed: 14769074]
- McCutchen D, Bell LC, France IM, Perfetti CA. Phoneme specific interference in reading: The tongue-twister effect revisited. *Reading Research Quarterly* 1991;26:87–103.
- McCutchen D, Perfetti CA. The visual tongue-twister effect: Phonological activation in silent reading. *Journal of Verbal Learning and Verbal Behavior* 1982;21:672–687.
- Perfetti CA, Bell L. Phonemic activation during the first 40 ms of word identification: Evidence from backward masking. *Journal of Memory and Language* 1991;30:473–485.
- Pexman PM, Lupker SJ, Jared D. Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 2001;27:139–156.

- Pollatsek A, Lesch M, Morris R, Rayner K. Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance* 1992;18:148–162. [PubMed: 1532185]
- Pollatsek A, Tan LH, Rayner K. The role of phonological codes in integrating information across saccadic eye movements in Chinese character identification. *Journal of Experimental Psychology: Human Perception and Performance* 2000;26:607–633. [PubMed: 10811166]
- Saito S, Baddeley A. Irrelevant sound disrupts speech production: Exploring the relationship between short-term memory and experimentally induced slips of the tongue. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* 2004;57:1309–1340.
- Stone GO, Van Orden GC. Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 1993;19:744–774. [PubMed: 8409857]
- Watkins KE, Strafella AP, Paus T. Modulation of motor excitability during speech perception: The role of broca's area. *Journal of Cognitive Neuroscience* 2003;16:978–987. [PubMed: 15298785]
- Wilson M. The case for sensorimotor coding in working memory. *Psychonomic Bulletin and Review* 2001;8:44–57. [PubMed: 11340866]
- Wilson SM, Saygin AP, Sereno MI, Iacoboni M. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience* 2004;7:701–702.

Table 1

Lexical decision times (measured in milliseconds) and error rates (in percentage) for Experiment 1 as a function of irrelevant spoken word type (ISW) for words and pseudowords. Standard errors of the mean are shown in parentheses.

	Identical ISW	Phon Similar ISW	Dissimilar ISW
LDT			
Words	686 (23)	732 (26)	710 (24)
Regular Pseudowords	799 (27)	787 (25)	781 (26)
Error			
Words	7 (1.2)	16 (1.5)	10 (1.4)
Regular Pseudowords	9 (1.4)	10 (1.4)	8 (1.3)

Table 2
Lexical decision times (in ms) and error rates (%) for Experiment 2 for words and pseudowords as a function of articulatory suppression type and irrelevant spoken word type. Standard error of the mean are presented in parentheses.

	<i>Without Articulatory Suppression</i>			<i>With Articulatory Suppression</i>		
	Identical ISW	Phon. Similar ISW	Dissimilar ISW	Identical ISW	Phon. Similar ISW	Dissimilar ISW
LDT						
Words	746 (32)	796 (33)	767 (29)	674 (26)	716 (31)	721 (30)
Pseudo-homophones	984 (32)	925 (72)	899 (35)	918 (34)	881 (33)	833 (29)
Regular Pseudowords	919 (47)	898 (43)	880 (44)	859 (37)	880 (44)	868 (41)
Error Rate						
Words	8 (2)	11 (2)	9 (2)	10 (2)	10 (2)	10 (2)
Pseudo-homophones	20 (3)	13 (4)	21 (3)	25 (4)	19 (4)	22 (3)
Regular Pseudowords	5 (2)	7 (3)	8 (2)	8 (2)	13 (3)	10 (2)