THE ROLE OF ORTHOGRAPHY IN SEGMENTATION OF SPEECH SOUNDS

Martha W. Burton & Donna Krebs Noble
Dept. of Neurology, University of Maryland School of Medicine, Baltimore, MD, USA
mburt001@umaryland.edu

ABSTRACT
Many pre-literate children, illiterate adults, and patients with acquired dyslexia have difficulty segmenting speech sounds despite intact or relatively spared spoken language abilities. Although it is unclear whether segmentation skills are a pre-requisite to or a consequence of reading ability, skilled readers may exploit orthographic knowledge during speech segmentation by invoking a strategy of visualizing the words to facilitate performance. To determine the role of orthographic knowledge in speech segmentation, 16 literate adults performed sound and letter judgment tasks on auditorily presented word pairs in which the consistency of orthographic and phonological information of the initial consonant was systematically varied (e.g., ‘card-cost,’ ‘card-kin,’ ‘guard-gem,’ ‘guard-jam’). Participants decided whether the first sound or letter was same or different. Task order was varied to explore possible strategic effects depending on the sequence of judgments. Participants were highly accurate on sound and letter judgments. For both tasks, responses to consistent phonological and orthographic pairs such as ‘card-cost’ were significantly more accurate than to conflicting pairs (e.g., ‘same’ sound pairs with conflicting spelling such as ‘card-kin,’ or ‘same’ letter pairs with conflicting initial phonemes such as ‘guard-gem’). In addition, ‘different’ judgments were slower on pairs with conflicting sound or spelling information in both tasks. Orthographic interference effects were found for the sound judgments even for participants who had not yet performed letter judgments. This effect could not be attributed to relying on a visualization strategy alone because of evidence of phonological interference effects during letter judgment, suggesting that orthographic and phonological knowledge in skilled readers may play an integrated role in segmentation of speech sounds.

INTRODUCTION
Typically, models of speech perception do not include a role for orthographic information (Klatt, 1979; Marslen-Wilson & Warren, 1994). However, lesion studies (e.g., Berndt, Haendiges, Mitchum, & Wayland, 1996) and priming studies (e.g., Jakimik, Cole, & Rudnicky, 1985) have suggested connections between spelling and speech perception. In particular, evidence has shown that many pre-literate children, illiterate adults, and patients with acquired dyslexia have difficulty segmenting speech sounds despite intact or relatively spared spoken language abilities (Berndt et al., 1996). Although it is unclear whether segmentation skills are a pre-requisite to or a consequence of reading ability, skilled readers may exploit orthographic knowledge during speech segmentation. Dijkstra et al. (1995) found phoneme monitoring response times were slowed for phonemes with secondary spellings compared to phonemes with a primary spelling. Such results have suggested that orthography affects the perception of spoken words (Ziegler & Ferrand, 1998).
One possibility is that listeners may invoke a strategy of visualizing the words to facilitate performance of these tasks. Alternatively, in literate readers, associations of orthographic and phonological information may be strongly learned associations that may automatically influence speech tasks, even those that do not require explicit access to such knowledge (e.g., phoneme discrimination, lexical decision).

To determine whether such orthographic effects are playing a role in a speech segmentation, we designed a phoneme discrimination experiment in which initial consonants with inconsistent mapping, such as ‘g,’ which can be pronounced as /g/ or /j/, were selected as initial consonants. We compared pairs of words in which sound and spelling of the initial phoneme had consistent mapping between members of the pair to those with mismatch of sound and spelling. If listeners do not access orthographic information during phoneme discrimination, no differences in sound judgments between pairs with matching and mismatching spelling should be expected. On the other hand, as has been suggested by previous studies, if orthographic knowledge plays a role in speech segmentation, pairs with conflicting sound and spelling information may degrade accuracy and/or slow response time.

**METHOD**

**Subjects**

Sixteen native speakers of American English (eight males and eight females between the ages of 18-46, mean = 30, SD = 6) with no known hearing impairments or learning disabilities participated in the study. The study was approved by the University of Maryland School of Medicine Institutional Review Board and all subjects provided written informed consent prior to participation.

**Stimuli**

The stimuli consisted of pairs of words with initial sounds that either had consistent or inconsistent matching of sound and spelling. There were a total of 100 experimental pairs of stimuli. For the pairs in which the initial phonemes were ‘same,’ half had matching initial spelling information (e.g., ‘card-cost,’ ‘giant-gel’) and the other half, mismatching orthographic information (e.g., ‘card-kin,’ ‘gem-jaw’). For the pairs with different initial phonemes, the phonological information was conflicting with the spelling judgment in half of the pairs (e.g., ‘ceiling-cuddle,’ ‘gem-game’), but not in the other half (e.g., ‘guard-jam’). In addition, 44 filler pairs with unambiguous sound and spelling of the initial phoneme served as distractors (e.g., ‘date-dime’, ‘date-tire’). Stimuli were divided into two lists with each condition equally represented in each list. Mean word frequency of the pairs was matched across conditions and lists. Although words occurred in more than one condition, no single word appeared more than three times in the experiment and no pairs of words were repeated.

Stimuli were recorded by a female native speaker of American English on DAT recorder (Sony TCD-D8) using a Sony ECM-MS957 microphone and edited on Pentium II PC running BLISS software at a sampling rate of 22050 Hz and 16-bit digitization. 50 ms of silence was inserted between the word pair members to separate the words. Mean pair duration was 1281 ms.
Procedures
Stimuli were presented on Pentium III laptop over Sony Headphones using E-Prime software (Psychology Software Tools, Pittsburgh, PA). Listeners performed two tasks, one in which they judged whether the first sound was same or different (sound judgment task) and the other in which they judged whether the first letter was same or different (letter judgment task). Each pair was presented with a stimulus onset asynchrony (SOA) of 5 sec. Half of the participants began the experiment by performing the sound judgment. Participants were then presented with a new list of the same type of pairs and judged whether the first letter was same or different. In the final set, they again rendered decisions on the first sound. A second group of subjects performed the two tasks in reverse order, starting with the letter judgment, followed by the sound judgment and then back to the letter judgment. The purpose of alternating the tasks was to determine if there were strategic changes in performance of the tasks within the same subjects depending on task order.

Subjects were instructed to perform the tasks as quickly and accurately as possible, indicating their response by pressing a button. Prior to the first run of each task, there was a set of 16 practice trials. Error rates and reaction times were measured for each participant.

RESULTS
Accuracy
Mean accuracy of the experimental pairs from each of the four conditions for sound and letter judgments across subjects are displayed in Table 1. These data were analyzed using a repeated measure ANOVA with the one between subjects factor, Task Order (Sound Judgment first (Group 1) versus Letter Judgment first (Group 2)), and three within subjects factors, Judgment (Sound versus Letter) X Condition (SP-SO (‘same phonology-same orthography’) versus SP-DO (‘same phonology-different orthography’) versus DPSO (‘different phonology-same phonology’) versus DP-DO (different phonology-different orthography)) X Session (1st presentation of task and 2nd presentation of task). There were two significant main effects. Specifically, participants were more accurate in the sound judgment task (96%) than in the letter judgment task (91%) [F(1, 14)=25.38, p < 0.01]. In addition, the ANOVA revealed a significant main effect of condition [F(3,42) =8.23, p <0.01]. Newman-Keuls post-hoc tests indicated that accuracy was lower on the conflicting pairs (i.e., SP-DO and DP-SO) than on the non-conflicting pairs (i.e., SP-SO, DP-DO) across both sound and letter judgments (p < .05). No main effect of group or any interaction with any of the other factors was found, indicating that neither task order nor session differentially affected accuracy of tasks and/or conditions.
Table 1. Mean % correct for sound and letter judgment tasks in each condition. Standard error of the mean is in parentheses. (Codes for conditions are: “s” = same, “d” = different, “p” = phonology, and “o” = orthography).

<table>
<thead>
<tr>
<th>Task</th>
<th>SP-SO ('card-cost')</th>
<th>SP-DO ('card-kin')</th>
<th>DP-SO ('guard-gem')</th>
<th>DP-DO ('guard-jam')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound Judgment</td>
<td>0.97 (0.01)</td>
<td>0.94 (0.02)</td>
<td>0.93 (0.03)</td>
<td>0.96 (0.02)</td>
</tr>
<tr>
<td>Letter Judgment</td>
<td>0.91 (0.02)</td>
<td>0.83 (0.04)</td>
<td>0.84 (0.04)</td>
<td>0.93 (0.02)</td>
</tr>
</tbody>
</table>

Response time
A four way repeated measures analysis of variance was performed on the reaction time data using the same factors as for accuracy. Figure 1 shows mean response times and standard error of the mean in each condition for each task. A significant main effect indicated that overall performance on the sound judgment task (1857 ms) was significantly faster than on the letter judgment task (2135 ms) \[F(1,14)=14.67, p < 0.01\]. In addition, there was an effect of session on reaction time with the second session (1925 ms) faster than the first session (2062 ms) \[F(1,14)=6.27, p < 0.05\]. A significant interaction between Judgment X Session indicated that the effect was larger in the letter judgment task \[F(1,14)=10.07, p < 0.01\]. Mean response time in the second session (2025 ms) was found to be significantly faster than in the first session (2247 ms) in the letter judgment task. However, in the phonology task, these differences were not significant \(p > .05\). The lack of a significant interaction between task order and any other factor, including session, suggests that strategies for performance of either the sound or letter judgment tasks were not affected by the initial or intervening task.

The ANOVA also revealed a significant main effect of condition \[F(3,42) = 6.88, p < 0.01\]. As in accuracy, conditions with confliction sound and spelling information were significantly slower than those did not have conflicts. In addition, the ANOVA revealed a Judgment X Condition interaction \[F(3,42)=5.11, p < 0.01\]. To explore this interaction further, we performed separate ANOVAs on the phonology and orthography judgment tasks. An ANOVA on the sound judgment task showed significant differences between the conditions \[F(3,42)=6.66, p < 0.01\]. Post-hoc tests showed mean response time to different phonology-same orthography (DP-SO) was significantly slower than both types of the non-conflicting pairs (i.e., SP-SO and DP-DO).

An ANOVA on the letter judgment task revealed significant differences between the conditions \[F(3,42)=5.61, p < 0.01\]. The mean response time to same phonology-different orthography (SP-DO) condition was significantly slower than either of the non-conflicting conditions (SP-SO, DP-DO) \(p < 0.05\) as well as different phonology-same orthography (DP-SO)(\(p < 0.05\)).
DISCUSSION

In summary, performance on the phonology task was more accurate and faster than the orthography task, indicating that the conversion of the phonemes to letters slowed listeners and reduced their accuracy. Because of the high accuracy on the conflicting pairs in the sound judgment task, it is evident that the participants were judging the initial sounds and not the spelling of those sounds. Participants were slowed on the sound judgment task when the words began with different sounds but had the same spelling (e.g., ‘guard-gem’), whereas in the letter judgment task, the words that began with the same sound but were spelled differently were the slowest (e.g., ‘card-kin’). Thus, in both tasks, mismatching sound and spelling information between members of the pairs slowed ‘different’ judgments.

These results are consistent with previous studies that have shown effects of spelling information on phonological tasks (Dijkstra et al., 1995). Although a possible strategy that listeners may rely on is visualizing the stimuli to perform their sound judgment, our results do no support such a view. If a visualization strategy were used, participants should have not shown effects of conflicting sound information in the letter judgment. Taken together, the results of the sound and letter judgment tasks indicate that in literate adults, phonological and orthographic information is integrated and is unaffected by strategies that may occur as a result of intervening tasks. Evidence of interference effects in listeners who are skilled readers suggest that for adults with reading skills that are not intact, performance on speech segmentation tasks may be further degraded if integration of sound and spelling is impaired.
ACKNOWLEDGEMENTS

The support of the NIH under Grant NIH DC R03-04004 to the University of Maryland Baltimore is gratefully acknowledged. The study was approved by the University of Maryland Baltimore Institutional Review Board and all participants provided written consent prior to participation.

REFERENCES