Neuropsychologia 111 (2018) 103–111

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Orthographic influence on spoken word identification: Behavioral and fMRI evidence

Christine Chiarello[a,*], Kenneth I. Vaden Jr.[b], Mark A. Eckert[b]

[a] Department of Psychology, University of California, Riverside, CA 92521, United States
[b] Medical University of South Carolina, United States

ARTICLE INFO

Keywords:
Auditory word identification
Cingulo-opercular system
Literacy
Lexical neighborhood effects
Orthography

ABSTRACT

The current study investigated behavioral and neuroimaging evidence for orthographic influences on auditory word identification. To assess such influences, the proportion of similar sounding words (i.e. phonological neighbors) that were also spelled similarly (i.e., orthographic neighbors) was computed for each auditorily presented word as the Orthographic-to-Phonological Overlap Ratio (OPOR). Speech intelligibility was manipulated by presenting monosyllabic words in multi-talker babble at two signal-to-noise ratios: + 3 and + 10 dB SNR. Identification rates were lower for high overlap words in the challenging + 3 dB SNR condition. In addition, BOLD contrast increased with OPOR at the more difficult SNR, and decreased with OPOR under more favorable SNR conditions. Both voxel-based and region of interest analyses demonstrated robust effects of OPOR in several cingulo-opercular regions. However, contrary to prior theoretical accounts, no task-related activity was observed in posterior regions associated with phonological or orthographic processing. We suggest that, when processing is difficult, orthographic-to-phonological feature overlap increases the availability of competing responses, which then requires additional support from domain general performance systems in order to produce a single response.

1. Introduction

Reading knowledge influences how speech is processed. Although spoken language is acquired prior to reading, literacy can affect some aspects of speech processing. For example, a word's orthography (spelling pattern) has an influence on how fast and accurately an auditory word is processed (Taft, 2011). Auditory lexical decisions are slower and less accurate for words with inconsistent sound-to-spelling mappings (Ziegler and Ferrand, 1998; Petrova et al., 2011). In addition, auditory rhyme judgments are facilitated when word pairs have similar (pie-tie), as compared to dissimilar (rye-tie), spellings (Seidenberg and Tanenhaus, 1979). Likewise, greater phonological priming of spoken words is obtained from words with more, than with less, spelling similarity (Chereau et al., 2007). Orthographic consistency effects have even been reported for task-irrelevant auditory words presented during a nonverbal noise detection task (Perre et al., 2011; see also Pattamadilok et al., 2014 for a related finding). Such findings support a beneficial effect of orthography on spoken word processing when word pairs have similar spellings or when potential competitors for a single spoken word are spelled similarly. It appears that orthographic representations are activated, perhaps automatically (Taft et al., 2008; Perre et al., 2011), during speech recognition.

Orthographic consistency studies generally examine the influence of words with rimes (vowel and any terminal consonants) that are spelled similarly or differently (e.g., Ziegler and Ferrand, 1998; Ziegler et al., 2004). Lexical neighborhood research explicitly examines how the processing of a target word is affected by the number of words with similar spellings or pronunciations, considering all of a word's segments.[1] Dense *phonological* neighborhoods have an inhibitory effect on auditory lexical decision (Vitevitch and Luce, 1999; Ziegler et al., 2003), object naming (Sadat et al., 2014), and shadowing (Dirks et al., 2001) tasks. However, the density of a word's *orthographic* neighborhood, that is the number of words sharing all but one letter with the target item, facilitates auditory word recognition in lexical decision and shadowing tasks (Ziegler et al., 2003). Such research examines the net size of a neighborhood rather than the extent to which orthographic and phonological neighborhoods overlap. Yet visual word recognition findings suggest that the number of such overlapping neighbors, rather than net neighborhood size, affects performance (Adelman and Brown, 2007). In the current investigation we take a similar approach to

* Corresponding author.
E-mail address: christine.chiarello@ucr.edu (C. Chiarello).

[1] Density calculations typically include words that differ from a target by a single phoneme, while consistency estimates only include words with identical rimes. The measures of interest in the current study are density estimates.

https://doi.org/10.1016/j.neuropsychologia.2018.01.032
Received 16 June 2017; Received in revised form 20 December 2017; Accepted 21 January 2018
Available online 31 January 2018

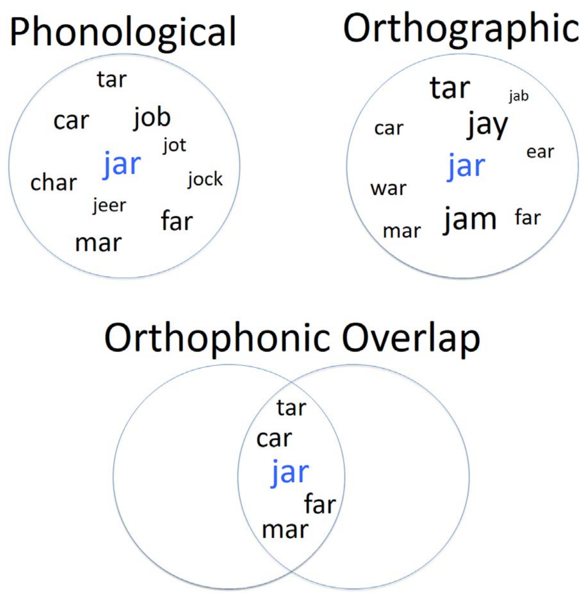

Fig. 1. Illustration of phonological and orthographic neighborhoods for 'jar' and the orthophonic overlap between them. Differences in word frequency approximated by font size. Not all neighbors are shown.

examine how orthography may influence aural word recognition accuracy by assessing the number of a spoken word's phonological neighbors that are also orthographic neighbors (orthophonic overlap - see Fig. 1).

Two theoretical approaches are consistent with a role for orthography in spoken word processing (Petrova et al., 2011; Taft, 2011). One approach (interactive activation) claims that sublexical units for orthography and phonology mutually influence each other and are jointly activated from speech. When such units converge on the same sublexical and lexical representations, processing is facilitated (Taft, 2011). However, for words with inconsistent spellings, competing inconsistent representations will impair processing. Another approach postulates that orthographic knowledge restructures how phonological information for words is stored, with finer-grained phonological representations for words with consistent spellings (Ziegler et al., 2008). According to this view, the locus of orthographic influence on speech identification is within the phonological system (Taft, 2011). Thus far, evidence from behavioral studies has been unable to distinguish between these possibilities (e.g., Ziegler et al., 2004; Petrova et al., 2011).

ERP data confirm an orthographic influence on speech recognition, with such effects detectable in waveforms by 300–400 ms post-onset (Perre and Ziegler, 2008; Perre et al., 2009a, 2009b). Across studies, these findings have been interpreted to support either the joint activation view (because orthographic and phonological priming effects differed in scalp topography, Perre et al., 2009a) or the phonological restructuring view (because the source of ERP spelling consistency effects was localized to left temporal-parietal sites associated with phonological processing, Perre et al., 2009b). Functional imaging during aural word recognition tasks could perhaps adjudicate between the theoretical approaches by examining the locus of brain activity associated with orthographic consistency/neighborhood size effects. The phonological restructuring view would predict predominant activity in regions associated with phonology, but not in areas associated with visual word processing. Only one prior fMRI study has explored this issue (Montant et al., 2011). No ventral occipital-temporal cortex (vOTC) effects were observed when comparing activation for inconsistent to consistent word spelling conditions during an auditory lexical decision task. There was elevated left frontal operculum activity that was interpreted as support for phonological restructuring of the "speech network."

Montant et al. (2011) frontal results, which included bilateral frontal operculum, anterior insula, and cingulate regions, can also be interpreted as involvement of a performance monitoring cingulo-opercular network that is typically engaged during challenging listening tasks (Eckert et al., 2016). Because this pattern of brain regions is recruited across verbal and nonverbal tasks (Dosenbach et al., 2008), it has also been interpreted as a domain general self-regulation system (Kelley et al., 2015) to guide behavior and optimize performance (Eckert et al., 2016). This performance monitoring and response selection perspective is generally consistent with the idea that orthographic neighborhood effects can occur at a decision or response selection stage of speech processing (Pattamadilok et al., 2007). Further support for this perspective may be seen in the lexical selection difficulty for orthographically inconsistent compared to consistent words that produces an effect at an FCz electrode location over medial-frontal cortex (Pattamadilok et al., 2009) where elevated theta power has been linked to an elevated threshold for response selection (Cavanagh et al., 2011).

Two other sources of data are relevant to the role of orthography in aural word recognition. Dehaene et al. (2010, 2015) have investigated how the acquisition of literacy affects the brain systems used to process spoken language. In individuals who became literate as adults, listening to spoken sentences or performing auditory lexical decisions was associated with increased activation bilaterally in the planum temporale (but with leftward asymmetry), as compared to matched illiterates (Dehaene et al., 2010). In addition, for auditory lexical decisions, literacy also was associated with increased activation in a region of the left inferior temporal cortex (the putative visual word form area), which the authors argue indicated recruitment of an orthographic code from speech. Similar findings were obtained when comparing child readers to age-matched non-readers during auditory sentence listening (Monzalvo and Dehaene-Lambertz, 2013). In that study, enhanced activation of the planum temporale was observed in 6-year-old readers compared to pre-readers. However, activation of the visual word form area was only increased in 9-year-old, compared to 6-year-old readers, implying that greater reading experience is required for literacy effects on speech in this region (Monzalvo and Dehaene-Lambertz, 2013). Dehaene and colleagues (2010, 2015) conclude from such studies that reading acquisition induces both restructuring of phonological representations in the planum temporale as well as activation of orthographic codes from speech.

Pattamadilok et al. (2010) applied transcranial magnetic stimulation (TMS) separately to the left supramarginal gyrus (SMG; a region associated with phonological processing) and to the left vOTC during auditory lexical decision. These stimulation sites were chosen based on independent experiments demonstrating disruption to phonological and orthographic processing, respectively. Stimulation to the left SMG, but not to the left vOTC, eliminated the response time benefit for orthographically consistent words. These authors propose that the SMG forms part of a "clean up" circuit that functions to aid perception in the face of noisy or distorted input. As reading is acquired, it provides an additional mapping onto phono-articulatory codes, strengthening their representations, and reducing the amount of clean up processing required to resolve the auditory input. Pattamodilok et al. (2010) conjecture that disruption of SMG clean up operations reduces or eliminates the processing advantage for orthographically consistent words. This is a novel interpretation, based on inhibitory transcranial stimulation of SMG, and is consistent with claims that orthographic influences may function to stabilize transient acoustic information (Ziegler et al., 2008). The Pattamodilok claim might predict that under noisy conditions, orthographic influences on speech identification should be enhanced and associated with increased "clean up" SMG activity when stimuli are presented in relatively poor signal-to-noise conditions. In the current study this prediction was tested by aurally presenting words at two SNR levels relative to a continuous multi-talker babble.