NeuroImage 89 (2014) 331–344

Contents lists available at ScienceDirect

NeuroImage

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

The similarity structure of distributed neural responses reveals the multiple representations of letters

David Rothlein [a,*], Brenda Rapp [a,b,c]

[a] Department of Cognitive Science, Johns Hopkins University, Baltimore MD 21218, USA
[b] Department of Neuroscience, Johns Hopkins University, Baltimore MD 21218, USA
[c] Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore MD 21218, USA

ARTICLE INFO

Article history:
Accepted 23 November 2013
Available online 7 December 2013

Keywords:
Letter representation
Abstract letter identity
Reading and spelling networks
Multivariate Pattern Analysis (MVPA)
Representational Similarity Analysis (RSA)
Visual Word Form Area (VWFA)

ABSTRACT

Most cognitive theories of reading and spelling posit modality-specific representations of letter shapes, spoken letter names, and motor plans as well as abstract, amodal letter representations that serve to unify the various modality-specific formats. However, fundamental questions remain regarding the very existence of abstract letter representations, the neuro-topography of the different types of letter representations, and the degree of cortical selectivity for orthographic information. We directly test quantitative models of the similarity/dissimilarity structure of distributed neural representations of letters using Multivariate Pattern Analysis–Representational Similarity Analysis (MVPA–RSA) searchlight methods to analyze the BOLD response recorded from single letter viewing. These analyses reveal a left hemisphere ventral temporal region selectively tuned to abstract letter representations as well as substrates tuned to modality-specific (visual, phonological and motoric) representations of letters. The approaches applied in this research address various shortcomings of previous studies that have investigated these questions and, therefore, the findings we report serve to advance our understanding of the nature and format of the representations that occur within the various sub-regions of the large-scale networks used in reading and spelling.

Introduction

Reading and spelling, as relatively recent additions to the human skill repertoire, presumably make use of evolutionarily older neural circuitry in order to represent the letter shapes, letter names and motor plans that are necessary for reading and writing. Consistent with this, extensive functional neuroimaging research has identified large-scale networks within visual, language and motor areas of the brain that are reliably recruited for reading and spelling. However, despite considerable consistency in the findings at a general level, there is far less consensus regarding how the component processes and representations are specifically instantiated within these networks. Various approaches have been taken for investigating these more detailed questions, each with specific strengths and weaknesses, as we discuss below. In the research we report on here, we directly address the question of the representational format of the neural codes used in reading by applying an MVPA–RSA searchlight analysis to fMRI data collected from subjects viewing single letters. We use this approach to identify the neuro-topographic distribution of the multiple codes of letters: abstract letter identities, visual letter shapes, letter names and motor programs for writing letter shapes. The findings of this research allow us to address long-standing cognitive science questions regarding the types of representations used in reading as well as neuroscience questions regarding the neural instantiation of these representations.

Multivariate Pattern Analysis (MVPA) (Kriegeskorte et al., 2008) is based on the premise that the informational content of neural representations is distributed across a population of neuronal units and, therefore, that stimuli that are representationally similar will generate similar response patterns across the neuronal units within a relevant brain region. The most common applications of MVPA involve the use of classification algorithms (Kriegeskorte, 2011) to determine if the patterns of responses within some brain region contain sufficient information to distinguish between two (or more) classes of stimuli (e.g., words vs. false fonts). However, instead of MVPA classification, in this work we use an MVPA Representational Similarity Analysis (MVPA–RSA) approach to model testing and comparison. This approach allows a comparison, within neural regions, of the *observed* similarity/dissimilarity structure of the voxel response patterns (e.g., to letter stimuli) with different quantitative models of the patterns that would be *predicted* if a region were sensitive to a specific type of representation (e.g., for letters: abstract, visual, phonological, or motoric). Furthermore, we specifically use a searchlight approach rather than the region of interest (ROI) approach that previously has been commonly used, even with

Abbreviations: ALI, Abstract letter identity; oRSM, Observed Representational Similarity Matrix; pRSM, Predicted Representational Similarity Matrix.

* Corresponding author at: Department of Cognitive Science, Johns Hopkins University, 237 Krieger Hall, 3400 North Charles Street, Baltimore, Maryland 21218, USA. Fax: +1 410 516 8020.

E-mail address: david.rothlein@gmail.com (D. Rothlein).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.11.054

RSA. We do so because the searchlight approach allows for model testing in a topographically neutral manner as a searchlight volume systematically examines large swathes of the brain.

Most functional neuroimaging research on orthographic processing has used words and other types of letter and letter-like strings as stimuli and has been directed at questions regarding the orthographic specificity of left ventral temporal-occipital cortex (e.g. Baker et al., 2007; Dehaene and Cohen, 2011), the unit-size of orthographic representations in this region (e.g., words or sub-lexical units) (Glezer et al., 2009; Nestor et al., 2013; Vinckier et al., 2007) or at questions concerning the format of orthographic representations within this area (Dehaene et al., 2001, 2004; Polk and Farah, 2002). In our work we do not address the "unit size" question, focusing instead on the question of representational format and content, without limiting ourselves to the ventral temporal-occipital region.

As we discuss below, the use of word and word-like stimuli raises specific interpretational challenges and so, to circumvent some of these, we use single letters as stimuli. A key advantage to single letter stimuli is that the multiple representations of letters are well-defined, dissociable and provide clear predictions for an RSA approach. Specifically, letters have characteristic visual shapes, spoken names and motor plans and these feature dimensions are dissociable in the sense that letters with similar visual shapes (e.g., A/R) may have different sounding names and motor plans, etc. In addition, many theories of reading and spelling posit abstract letter representations (ALIs) (Brunsdon et al., 2006; Jackson and Coltheart, 2001) that serve to unify and mediate between the different cases, fonts and modality-specific formats, such that E, e, 𝔈 and/i/all correspond to precisely the same abstract representation (Fig. 1). For these reasons, an MVPA–RSA investigation of neural responses to single letters is well-suited to addressing questions of representational format. The prediction is that if there are regions specifically tuned to abstract letter identities, letter shapes, names or motor plans they should produce similar neural responses for letters that are similar along one of these dimensions but not for letters that are similar along others. For example, in an area that specifically encodes visual letter shapes, the pattern of responses across voxels should be correlated when participants view letters with similar shapes (A/R) but not when they view visually dissimilar letters (A/S), nor when they view letters with only phonologically similar names (B/P) or similar motor plans (T/L). The same logic extends to the other modality-specific representational types. With respect to abstract letter identities (ALIs), the prediction is that neural substrates encoding these representations should respond similarly to letters that have the same identity despite differing in case and visual appearance (A/a). The further key prediction is that substrates that selectively encode ALIs should be insensitive to similarities between letters in terms of their visual–spatial, letter-name or motoric features.

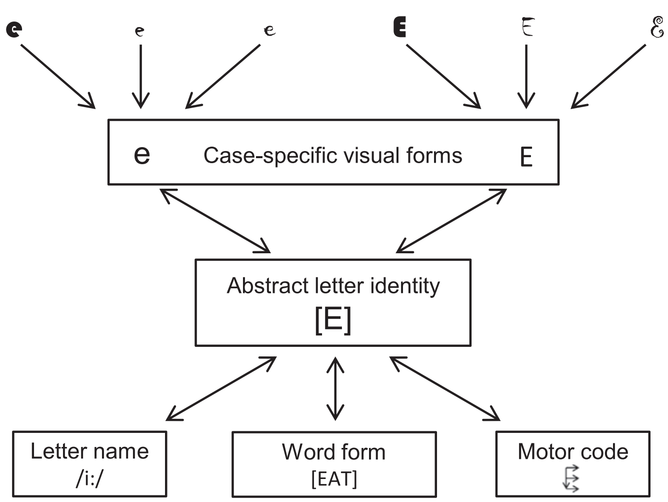

Fig. 1. The multiple formats of letter representation. Visual–spatial letter representations are case-specific representations that are invariant across different fonts. Letter-name representations correspond to the names of letters, while Motoric representations correspond to the basic motor sequences required to produce letter shapes. Abstract Letter Identities (ALIs) are amodal, abstract representations that lack visual form (they are font and case invariant), phonological content, or motor features. ALIs mediate translation between modality-specific formats. Word reading is based on ALIs that are accessed in response to processing specific visual letter shapes and are then used to search memory for the stored orthographic representations of familiar word forms.

Theories of reading (e.g., Grainger et al., 2008) often distinguish between low-level representations of visual features of letters, high-level representations of letter shapes and abstract letter identities (ALIs) (Fig. 1). Low-level visual-feature representations correspond to those involved in visual processing more generally and would include representations computed in primary and early visual areas. At this level, the same letter in a different font (e.g., A/A) would be represented differently. In contrast, at the level of visual-shape representations, the underlying shape/geometry of a letter is represented, in a manner comparable to what is sometimes referred to in the visual object processing literature as a "structural description" (Miozzo and Caramazza, 1998). At this level, letters in different fonts (A/A) would share a representation but different allographs of a letter (A/a), would not. A further distinction is made by reading theories that assume that letter-shape representations are recoded into abstract letter identities (Jackson and Coltheart, 2001) that are used to search memory for stored orthographic representations of familiar word forms. Because ALIs are abstract (font and case invariant) *A* and *a* would be represented in the same way. ALIs allow readers to easily recognize words in unfamiliar fonts or case (𝒫eAc𝓗). Of these three letter representation types, ALIs are the most controversial, both in terms of their existence and also with regard to their neural instantiation. Alternatives to ALI-mediated views posit that reading is based on either visual representations of letters alone or on visual exemplars of previously experienced words and letters (Tenpenny, 1995; for example, see Plaut and Behrmann (2011) for a model of letter representation that does not include ALI's). In fact, the contrast between the ALI vs. visually mediated views of reading is a clear example of the larger, long-standing and contentious debate between the abstractionist and grounded cognition (or semantic and episodic) views of human knowledge representation (Barsalou, 2008; Tulving, 1983).

Considerable behavioral (Besner et al., 1984; Kinoshita and Kaplan, 2008), neuropsychological (Coltheart, 1981) and neuroimaging (Dehaene et al., 2004; Polk and Farah, 2002) evidence has been put forward in support of ALIs. With regard to neural substrates, neuroimaging research has generally localized ALIs to the posterior, inferior temporal lobe (Dehaene et al., 2004). This is consistent with the mid-fusiform localization of the Visual Word Form Area (VWFA) assumed by many to be an orthographic processing area critical for word reading (Cohen et al., 2000; Tsapkini and Rapp, 2010). However, the attribution of ALIs to this brain area remains highly debated (Barton et al., 2010a, 2010b; Burgund and Edwards, 2008; Price and Devlin, 2003; Wong et al., 2009). Critically, as we review next, previous studies arguing for ALIs have not controlled for the possibility that effects attributed to ALIs might instead originate from modality-specific (visual, phonological, motor) or semantic representations of the word or letter stimuli used in these studies.

Some studies examining the nature of orthographic representations have reported similar neural responses or priming effects for orthographic stimuli presented in different fonts (Gauthier et al., 2000; Nestor et al., 2013; Qiao et al., 2010). However, while similar responses to different-font stimuli indicate that the recruited neural representations are indeed more abstract than low-level visual feature representations, the similar responses may have originated from visual letter-shape representations rather than from ALI representations. Thus, the finding of similar cross-font responses does not necessarily implicate ALIs. Other studies have specifically manipulated letter case in order to examine issues of representational format. For example, some studies have shown that activity in the left mid-fusiform was comparable for uniform-case and mixed-case words and pseudowords (APPLE/aPpLe)