Journal of Experimental Child Psychology 116 (2013) 37–44

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Child Psychology

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

Contextual diversity is a main determinant of word identification times in young readers

Manuel Perea [a,*], Ana Paula Soares [b], Montserrat Comesaña [b]

[a] ERI – Lectura and Departamento de Metodología, Facultad de Psicología, Universitat de València, 46010 Valencia, Spain
[b] Human Cognition Lab, CIPsi, School of Psychology, University of Minho, 4710-057 Braga, Portugal

ARTICLE INFO

Article history:
Received 12 September 2012
Revised 28 October 2012
Available online 29 January 2013

Keywords:
Word identification
Word frequency
Lexical decision
Developing readers
Reading
Computational models

ABSTRACT

Recent research with college-aged skilled readers by Adelman and colleagues revealed that contextual diversity (i.e., the number of contexts in which a word appears) is a more critical determinant of visual word recognition than mere repeated exposure (i.e., word frequency) (*Psychological Science*, 2006, Vol. 17, pp. 814–823). Given that contextual diversity has been claimed to be a relevant factor to word acquisition in developing readers, the effects of contextual diversity should also be a main determinant of word identification times in developing readers. A lexical decision experiment was conducted to examine the effects of contextual diversity and word frequency in young readers (children in fourth grade). Results revealed a sizable effect of contextual diversity, but not of word frequency, thereby generalizing Adelman and colleagues' data to a child population. These findings call for the implementation of dynamic developmental models of visual word recognition that go beyond a learning rule by mere exposure.

Introduction

Word identification times in laboratory tasks and eye fixation times during normal reading are shorter when reading a word that occurs frequently in print such as *heart* than when reading an infrequent word such as *elbow*. This is the case both in adult readers (e.g., see Balota et al., 2007; Inhoff & Rayner, 1986) and in young readers (e.g., Hyönä & Olson, 1995; Moret-Tatay & Perea, 2011).

Unsurprisingly, word frequency plays a central role in all current computational models of visual word recognition and reading. For instance, in the family of localist activation-based models, the

* Corresponding author.
E-mail address: mperea@valencia.edu (M. Perea).

0022-0965/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jecp.2012.10.014

resting level of activation of a given word unit depends on its printed word frequency (e.g., spatial coding model: Davis, 2010; dual-route cascaded model: Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; interactive activation model: McClelland & Rumelhart, 1981; multiple read-out model: Grainger & Jacobs, 1996). Likewise, computational models of reading such as the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998) and the SWIFT model (Engbert, Nuthmann, Richter, & Kliegl, 2005) employ a similar mechanism of word frequency during the initial "lexical access" stage.

One obvious shortcoming of these "static" models is that they assume a well-formed and unalterable set of parameters (as in an adult skilled reader) in which word frequency is included as a critical element (see Baayen, 2010, for criticism of these models). Although there is the implicit assumption that the frequency effect has its origin in the repetition effect, these models provide no information about the underlying nature (or locus) of the word frequency effect. Thus, one obvious question is how repeated presentations affect a word's accessibility during the course of word learning in an immature word recognition system (see Acha & Perea, 2008; Wang, Castles, Nickels, & Nation, 2011, for developmental analyses of orthographic effects in visual word recognition).

Is word accessibility merely driven by a count of past presentations (i.e., word frequency)? Recently, a number of theorists have claimed, using large databases collected from college-aged students, that the main determinant of word identification times is not "pure repetition" per se but rather the number of contexts in which a given word occurs (Adelman, Brown, & Quesada, 2006; see also Baayen, 2010; Brysbaert & New, 2009, & Keuleers, Brysbaert, & New, 2010, for similar conclusions). Clearly, words that appear in many contexts tend to be quite frequent in print, and this explains the previous reports of a word frequency effect; they would be obtained because of a confound with the number of contexts in which a word occurs. This latter factor has been named *contextual diversity* and has been defined as "the proportion of texts in which a given word occurs" (Adelman et al., 2006). Adelman and colleagues (2006) indicated that if a given word appears in very different contexts, the likelihood of that word appearing in other contexts increases, and this may facilitate the relative accessibility of these words in the "mental lexicon".

Importantly, contextual diversity has also been claimed to be a relevant factor to word acquisition in developing readers (e.g., see Hills, Maouene, Riordan, & Smith, 2010). Thus, if words that are initially learned in many contexts have more accessible and enriched internal representations than words that are learned in a limited number of contexts, then the effects of contextual diversity should be particularly clear for developing readers. (Note that children's word learning mechanisms may be even more important for word recognition than in the case of the more "stable" word representations in skilled adult readers; see Ratcliff, Love, Thompson, & Opfer, 2012, for evidence of less efficient word processing in children than in college-aged students.)

In the current experiment, we examined the role of contextual diversity and word frequency in word identification times in young readers (children in fourth grade). Unsurprisingly, it was not possible to obtain enough low-frequency words with high contextual diversity, and this latter condition was not included in the experiment. The hypotheses are clear. If mere repeated exposure (and not context diversity) is the driving force in lexical access (as stated in currently implemented computational models), then words that occur frequently in print—while keeping contextual diversity controlled—should produce faster word identification times than words that occur least frequently (i.e., a word frequency effect). Alternatively, if context diversity (and not word frequency) is the driving force in lexical access (as claimed by Adelman and others), then words that occur in many contexts in print—while keeping word frequency controlled—should produce faster word identification times than words that occur in few contexts. If this latter hypothesis is confirmed, then a new conceptualization of models of visual word recognition would be required. This applies to more than just the static activation-based models cited above. Connectionist models (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996) are dynamic models that include a learning rule (i.e., word learning is possible). However, as noted by Baayen (2010), this learning rule relies on repeated presentations of the words rather than on contextual diversity (i.e., the PDP [parallel distributed processing] model cannot explain the contextual diversity effects reported by Adelman et al., 2006).

Contextual diversity and word frequency were obtained from the ESCOLEX database (Soares et al., 2012). ESCOLEX is a children's lexical database with grade-level word frequency statistics for 6- to 11-year-olds (first to sixth grades) for a total of approximately 48,500 word forms computed from a