Cognitive Psychology 70 (2014) 1–30

Contents lists available at ScienceDirect

Cognitive Psychology

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

How lexical is the lexicon? Evidence for integrated auditory memory representations

April Pufahl [a,1], Arthur G. Samuel [a,b,c,*,2]

[a] Department of Psychology, Stony Brook University, 100 Nicholls Road, Stony Brook, NY 11794-2500, USA
[b] IKERBASQUE Basque Foundation for Science, Alameda Urquijo, 36-5, Plaza Bizkaia, 48011 Bilbao, Bizkaia, Spain
[c] Basque Center on Cognition Brain and Language, Paseo Mikeletegi 69, 2nd Floor, 20009 Donostia-San Sebastián, Gipuzkoa, Spain

ARTICLE INFO

Article history:
Accepted 6 January 2014

Keywords:
Implicit memory
Indexical effects
Mental lexicon
Priming
Specificity effects
Spoken word recognition

ABSTRACT

Previous research has shown that lexical representations must include not only linguistic information (what word was said), but also indexical information (how it was said, and by whom). The present work demonstrates that even this expansion is not sufficient. Seemingly irrelevant information, such as an unattended background sound, is retained in memory and can facilitate subsequent speech perception. We presented participants with spoken words paired with environmental sounds (e.g., a phone ringing), and had them make an "animate/inanimate" decision for each word. Later performance identifying filtered versions of the words was impaired to a similar degree if the voice changed or if the environmental sound changed. Moreover, when quite dissimilar words were used at exposure and test, we observed the same result when we reversed the roles of the words and the environmental sounds. The experiments also demonstrated limits to these effects, with no benefit from repetition. Theoretically, our results support two alternative possibilities: (1) Lexical representations are memory representations, and are not walled off from those for other sounds. Indexical effects reflect simply one type of co-occurrence that is incorporated into such representations. (2) The existing literature on indexical effects does not actually bear on lexical representations – voice changes, like environmental sounds heard with a word, produce implicit memory effects that are not tied to the lexicon. We discuss the evidence and implications of these two theoretical alternatives.

* Corresponding author at: Department of Psychology, 100 Nicholls Road, Stony Brook University, Stony Brook, NY 11794-2500, USA. Fax: +1 631 632 7876.
E-mail addresses: april.pufahl@stonybrook.edu (A. Pufahl), arthur.samuel@stonybrook.edu, a.samuel@bcbl.eu (A.G. Samuel).
[1] Fax: +1 631 632 7876.
[2] Fax: +34 943 309 052.

0010-0285/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cogpsych.2014.01.001

1. Introduction

Fundamentally, the mental lexicon is a memory system: It is the place where language and memory meet. Most models of spoken word recognition (e.g., TRACE: McClelland & Elman, 1986; shortlist: Norris, 1994; PARSYN: Luce, Goldinger, Auer, & Vitevitch, 2000; distributed cohort model: Gaskell & Marslen-Wilson, 1997, 1999, 2002) assume the incoming speech signal is mapped onto abstract linguistic representations. As such, in these models, the input codes for lexical representations include only abstract phonological features that differentiate between words. One major challenge to this assumption comes from empirical evidence that speech recognition is sensitive to changes in surface characteristics such as the voice of the speaker – a set of properties that collectively constitute "indexical" information. These specificity effects have led to an expansion of the mental lexicon to include episodic features reflecting this indexical variation (Goldinger, 1996, 1998, 2007; Johnson, 1997, 2005, 2006; Palmeri, Goldinger, & Pisoni, 1993; Pierrehumbert, 2001; Sheffert, 1998). Other models have retained abstract linguistic representations but also included probabilistic information about their occurrence that can be altered based on input by a given speaker (Clayards, Tanenhaus, Aslin, & Jacobs, 2008; Norris & McQueen, 2008).

In the present study, we ask whether the expansion of the mental lexicon to include both linguistic and indexical information is sufficient. Arguably, voices are an important source of variation in comprehending spoken language, so the inclusion of indexical information in the mental lexicon may serve a pragmatic purpose. On the other hand, it is impossible to hear a word without also hearing the voice speaking it, so the inclusion of indexical information in the mental lexicon could simply be due to its co-occurrence with linguistic information. From this perspective, the indexical properties added to some speech recognition theories are not necessarily indexical *per se*, but simply properties that happen to be co-present with the linguistic information. To test this possibility, we compared the co-occurrence of words and voices to the co-occurrence of words and irrelevant environmental sounds. Given that speech and non-speech sounds are frequently encountered simultaneously, how does the system treat additional variation from this co-occurring non-speech? Do listeners discard variability in the incoming auditory signal that comes from non-human sources when attending to speech, or does this variability, like that from voices, persist in memory?

1.1. Talker variability in speech perception

Previous research has shown that listeners retain speaker-specific auditory details in memory, and that these memories help facilitate future understanding of previously encountered speakers (for a review, see Luce & McLennan, 2005). These indexical effects refer to any performance advantage (e.g., improved accuracy or response time) for tokens repeated in the same voice (or with similar properties) over a different voice.

In a typical indexical study, participants first perform a task to encode the stimuli into memory. After some delay, they then complete a memory test with stimuli repeated in the same voice (or with similar properties) or in a different voice. Encoding tasks have varied in terms of depth of processing, such as classifying words according to the speaker's gender (shallow), initial phoneme (moderate), and syntactic class (deep; Goldinger, 1996). Other encoding tasks have drawn attention to the voice by requiring participants to identify the speaker (Allen & Miller, 2004; Goldinger, 1996; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994) or to rate the pitch or clarity of pronunciation (Church & Schacter, 1994; Schacter & Church, 1992). These tasks contrast with others that do not require processing of voice characteristics, such as making a "word/non-word" lexical decision (Luce & Lyons, 1998), identifying the category to which a word belongs (Schacter & Church, 1992), or counting the number of meanings for a word (Church & Schacter, 1994; Schacter & Church, 1992). Overall, indexical effects appear to be relatively insensitive to task changes at encoding, as evidenced by a performance advantage for same-voice over different-voice items across levels of processing (Goldinger, 1996) and across tasks that do and do not draw attention to voice characteristics (Schacter & Church, 1992).