Journal of Memory and Language 93 (2017) 104–134

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Spoken-word recognition in 2-year-olds: The tug of war between phonological and semantic activation

Janette Chow [a,*], Anne Aimola Davies [a,b], Kim Plunkett [a]

[a] Department of Experimental Psychology, University of Oxford, United Kingdom
[b] Research School of Psychology, Australian National University, Australia

ARTICLE INFO

Article history:
Received 1 August 2015
revision received 8 August 2016
Available online 11 October 2016

Keywords:
Visual world paradigm
Lexical access
Phonology
Semantics
Eye tracking
Toddler

ABSTRACT

Previous studies demonstrate that while toddlers can match words with their referent before the age of one, they only begin to extract phonologically- and semantically-related information from speech later in the second year. However, the order and manner in which this information is extracted remains unresolved. In two experiments, we adapted the adult four-picture Visual World Paradigm (VWP) (Huettig & McQueen, 2007) for toddler testing: toddlers hear a spoken word and see pictures that are phonologically-related, semantically-related or unrelated to the spoken-word. We demonstrate that, similar to adults, 24- to 30-month-old toddlers attended to a phonological match faster than a semantic match, and that these differences hold irrespective of whether the semantic match is taxonomic or thematic in character. Our findings suggest that language shapes toddlers' cognition by biasing their selective attention to relevant information in the visual world, which is likely to enhance the efficiency of mental activities, such as learning and making analogical inferences.

Introduction

Toddlers can match words with their referents before the age of one. Do toddlers simply associate a spoken label directly with its referent, much as a pigeon learns that a light signals reward, or do they also access phonologically- and semantically-related information from the lexical-semantic system in a cascaded fashion as has been observed with adults? Numerous studies have demonstrated that toddlers, by the age of two, extract phonologically- and semantically-related information from speech. However, the order and manner in which this information is extracted remains unresolved. This study provides the missing piece of evidence by examining the time course of the online mental processes involved in spoken-word recognition in 24- to 30-month-old toddlers by adapting the adult Visual World Paradigm (VWP) for toddler testing. Furthermore, this study attempts to provide insights into the manner in which spoken-word recognition drives visual attention during toddlerhood, a critical period of language acquisition. In particular, how does the encoding of phonological and semantic relations affect the way toddlers direct their attention to objects in the world?

Over the past 20 years, the VWP has become a popular tool for the study of language processing in adults (see Huettig, Rommers, & Meyer, 2011, for a detailed review). One line of VWP research focuses on measuring the manner in which lexical-semantic representations are activated during spoken-word recognition. Cooper (1974) first demonstrated experimentally that adults would spontaneously and automatically fixate an object that was named or relevant to what was just said. In his studies, participants were presented with a grid of nine pictures, and heard a story. When participants heard the word 'lion',

* Corresponding author.
E-mail address: janette.chow@psy.ox.ac.uk (J. Chow).

http://dx.doi.org/10.1016/j.jml.2016.08.004
0749-596X/

their eyes were drawn to the picture of the lion. Upon hearing the word 'Africa', they were also more likely to fixate the pictures of lion, zebra and snake. These findings demonstrate that visual attention can be driven by the matching of auditory and visual stimuli, even when there is no explicit task or requirement for an overt response. The fact that participants fixated the related but unnamed pictures suggests that target-related representations were activated and retrieved from the lexical-semantic system, and matched with the current visual input, which in turn led to increased visual attention. Cooper's (1974) study provided concrete evidence that language mediates the listener's eye-movements, and that language-mediated eye-movements reflect mental activities of language processing.

Phonological representations

The Cohort model of word recognition (Marslen-Wilson, 1987) proposes that the initial sounds of a word activate a set of candidate words in the lexicon that begin with the same phonemes. As the spoken word unfolds, candidate words that no longer match the speech input are eliminated until only one word remains. For example, upon hearing the first sound /d/ of the word 'dog', candidate words, such as 'dog', 'digger' and 'dress', are activated. Once the spoken vowel is processed, 'digger' and 'dress' can be discarded. The Cohort model assumes that lexical access is a strictly bottom-up process; top-down effects, such as context, only influence word selection among the candidate words. The TRACE model (McClelland & Elman, 1986) further embodies the proposal that spoken word recognition is not only incremental but also continuous, and therefore, words that match the sensory input in the middle or at the end of the signal, such as rhyme words, will also be activated (e.g., 'dog' will activate 'log'). Using the VWP, Allopenna, Magnuson, and Tanenhaus (1998) presented participants with a four-picture display including a target, cohort distractor (same onset consonant and first vowel as target), rhyme distractor (rhymes with target) and unrelated distractor. Upon hearing the target word, participants not only fixated the target picture, but also fixated the cohort distractor. Later in the time course, participants also fixated the rhyme distractor more than the unrelated distractor. Their findings provide strong evidence that spoken-word recognition involves continuous incremental mapping of speech sounds and the automatic activation and elimination of competing phonological representations.

Semantic representations

VWP studies have also provided evidence that spoken-word recognition leads to the automatic activation of semantically-related representations. Taxonomic and thematic relations are the two most studied types of semantic relations. Taxonomic relations refer to the classification or hierarchy of items. Items that are taxonomically-related tend to resemble each other, sharing perceptual, biological or functional properties (e.g., Medin & Smith, 1984; Smith & Medin, 1981). Taxonomic categories can be formed at different levels in the hierarchy, such as basic, subordinate and superordinate level categories. For example, labrador and collie belong to the category of dog; dog and horse belong to the category of mammals; mammals and birds belong to the category of animal. In contrast, thematic relations refer to the grouping of items co-occurring in the same space and time, often based on memory and personal experience. Rather than resembling each other, thematically-related items tend to complement each other. Amongst the most studied types of thematic relations are functional relations. These include at least two sub-types: instrument and script (Moss, Ostrin, Tyler, & Marslen-Wilson, 1995). The instrument relation describes word pairs in which one performs some action on the other. For example, shampoo is used on hair. The script relation describes the relationship between a context and items that occur or appear in the same time and space. For example, one finds sheep on a farm. Other commonly studied thematic relations include part-whole (e.g., coin is made of copper), causal-relation (e.g., wind propels a sail) and temporal-relation (e.g., breakfast takes place before lunch). Often, thematically-related items are related in more than one way (Lin & Murphy, 2001). For example, chalk and blackboard are both instrument- and script-related, because one uses chalk to write on the blackboard, and you often find a chalk resting on the ledge of a blackboard.

Studies using the VWP have shown that upon hearing a target word (e.g., 'dog'), the listener is more likely to fixate a taxonomically-related picture (e.g., pig) (Huettig & Altmann, 2005) or a thematically-related picture (e.g., bone) (Mirman & Graziano, 2012; Yee, Overton, & Thompson-Schill, 2009) than an unrelated distractor. These findings are consistent with localist models of semantic memory (e.g., Collins & Loftus, 1975) in which concepts are represented by nodes in a semantic network, and that concepts having a similar meaning share connections with each other within the network. When spoken-word or object recognition takes place, the activation of the referent node spreads to related nodes (i.e., semantically-related concepts), thereby underpinning language-mediated eye-movements in a VWP task. However, these results can also be taken as evidence for distributed models (e.g., Cree & McRae, 2003; Moss, Hare, Day, & Tyler, 1994; Tyler & Moss, 2001). Distributed or featured-based models argue that objects are categorised based on the degree of feature correlation or overlap. Correlation between taxonomically-related objects is achieved through the high number of shared-features (e.g., head and body in mammals). In contrast, the objects' distinctive, uncorrelated features (e.g., zebras have stripes and deers have antlers) allow discrimination within a category (Cree & McRae, 2003; Tyler & Moss, 2001). Thematically-related objects can be correlated in terms of their co-occurrence in episodic memories or the co-occurrence of lexical forms (e.g., 'shampoo' and 'hair') in spoken utterances or written text (Moss et al., 1994; Plaut & Booth, 2000). In the VWP, fixation of taxonomically- and thematically-related pictures is motivated by the activation of shared features and co-occurrence associations with the target.