Cognition 132 (2014) 429–436

Contents lists available at ScienceDirect

Cognition

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

Distributional structure in language: Contributions to noun–verb difficulty differences in infant word recognition

Jon A. Willits [a,*], Mark S. Seidenberg [b], Jenny R. Saffran [b]

[a] Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St., Bloomington, IN 47405, United States
[b] Department of Psychology, University of Wisconsin, 1202 W. Johnson Street, Madison, WI 53706, United States

ARTICLE INFO

Article history:
Received 2 December 2009
Revised 8 May 2014
Accepted 9 May 2014

Keywords:
Word recognition
Language acquisition
Verb learning
Statistical learning

ABSTRACT

What makes some words easy for infants to recognize, and other words difficult? We addressed this issue in the context of prior results suggesting that infants have difficulty recognizing verbs relative to nouns. In this work, we highlight the role played by the distributional contexts in which nouns and verbs occur. Distributional statistics predict that English nouns should generally be easier to recognize than verbs in fluent speech. However, there are situations in which distributional statistics provide similar support for verbs. The statistics for verbs that occur with the English morpheme *–ing*, for example, should facilitate verb recognition. In two experiments with 7.5- and 9.5-month-old infants, we tested the importance of distributional statistics for word recognition by varying the frequency of the contextual frames in which verbs occur. The results support the conclusion that distributional statistics are utilized by infant language learners and contribute to noun–verb differences in word recognition.

© 2014 Published by Elsevier B.V.

1. Introduction

How do infants come to recognize words in fluent speech, and why do they find some words easier to recognize than others? This question is a complex one, and is difficult to answer. But the answer is important, because word recognition is a critically important process with consequences for many downstream language processes. If infants do not recognize a familiar word in fluent speech, their ability to match its sounds to its referent will be impaired (Graf Estes, Evans, Alibali, & Saffran, 2007; Hay, Pelucchi, Graf Estes, & Saffran, 2011). Familiar words that go unrecognized will not be available as a cue to the boundaries of novel words in fluent speech (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Mersad & Nazzi, 2012). Further, an infant who is not recognizing words in fluent speech is missing out on a wealth of distributional information about the word's co-occurrences with other words (Fernald, Perfors, & Marchman, 2006; Swingley, Pinto, & Fernald, 1999). Infants' ability to quickly and efficiently recognize words in context is a strong predictor of vocabulary growth and other aspects of language processing (Fernald & Marchman, 2012; Marchman, Fernald, & Hurtado, 2010). These facts demonstrate the importance of a thorough understanding of why infants can recognize some words easily, but have difficulty with others.

The goal of the current research was to further investigate this process by examining the role that words' distributional contexts might play in facilitating infant word recognition. As a means to that end, we studied a specific case in which infants have word recognition challenges: recognizing verbs (relative to nouns). A large body of research, involving many languages and methodologies, indicates that children learn about nouns more easily than verbs (see Golinkoff & Hirsh-Pasek, 2006, for a review). It has been suggested that this noun advantage may begin

* Corresponding author. Tel.: +1 6086288066.
E-mail address: jon.willits@gmail.com (J.A. Willits).

http://dx.doi.org/10.1016/j.cognition.2014.05.004
0010-0277/© 2014 Published by Elsevier B.V.

at the earliest stages of language acquisition, with infant word recognition. For example, whereas native-language nouns can be reliably recognized in continuous speech by 6.0–7.5 month olds (Bortfeld et al., 2005; Jusczyk & Aslin, 1995), similar performance with verbs does not appear to emerge until 11.0–13.5 months of age (Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk, 2005; Shi & Lepage, 2008).

A number of factors are known to be important contributors to word recognition difficulty. One of the most critical factors is a word's acoustical properties and sound structure. For example, infants have a much easier time recognizing words that follow their native language's phonotactic regularities, such as being consonant-initial in English (Jusczyk & Aslin, 1995; Nazzi et al., 2005). Infants also have an easier time recognizing sound sequences that follow their native language's typical stress pattern (Echols, Crowhurst, & Childers, 1997; Houston, Santelmann, & Jusczyk, 2004; Jusczyk, Houston, & Newsome, 1999; Morgan & Saffran, 1995; Nazzi et al., 2005).

However, there is also variation that cannot be explained by phonotactic and prosodic structure. For example, Nazzi et al. (2005) carefully manipulated verbs' phonotactics and prosody, and found big effects: infants recognized the "easy" verbs (consonant-initial verbs with strong–weak stress) at 13.5 months and did not recognize "hard" verbs (vowel-initial verbs with weak-strong stress) until 16.0 months. But 10.5-month-old infants failed to recognize even the "easy" verbs, a marked contrast with studies showing that infants as young as 6.0–7.5-month-olds recognize nouns in fluent speech (Bortfeld et al., 2005; Jusczyk & Aslin, 1995). Thus despite a considerable amount of focus on infants' word recognition abilities, and the identification of phonotactic and prosodic factors that contribute to noun–verb differences, a thorough understanding of this difference eludes us.

An additional factor, and one that might help explain the noun–verb gap, is the infant's familiarity with the immediate contexts in which nouns and verbs occur. Natural language statistics predict what words will be easy for adults to recognize (Duffy, Morris, & Rayner, 1988), comprehend (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Gennari & MacDonald, 2009; Trueswell, Tanenhaus, & Kello, 1993), and produce (Haskell, Thornton, & MacDonald, 2010). And it is known that very young language learners engage in statistical learning, which occurs over various elements in the linguistic input as well as the environments in which utterances occur (Saffran, Aslin, & Newport, 1996; Smith, Yu, & Pereira, 2011). It has also been established that nouns and verbs differ systematically with respect to language statistics that affect language processing. In general, these differences favor nouns over verbs (Hills, 2013; St. Clair, Monaghan, & Christiansen, 2010; Willits, Seidenberg, & Saffran, 2009), consistent with the overall noun advantage observed in behavioral studies. We present new behavioral and quantitative evidence supporting this explanation, complementing previous findings.

However, the distributional context explanation makes a further prediction. If these distributional statistics are an important factor in word recognition, then in situations where verbs *are* more similar to nouns with respect to relevant distributional statistics, they should be as easy to recognize as nouns. English affords a way to test this hypothesis. In English, distributional statistics generally favor nouns over verbs, with more frequently and consistently occurring distributional frames. For example, nouns most frequent distributional collocation (*the*) has a considerably higher co-occurrence probability with nouns ($p = 0.192$) than verbs most frequent collocation (*you*, $p = 0.117$).[1] Nouns on average also co-occur with a smaller subset of words than verbs – giving them more consistent contextual collocation cues [$\overline{X} = 3.55\%$ (SE = 0.23%) for nouns, compared to $\overline{X} = 4.50\%$ (SE = 0.40%) for verbs, $t(299) = 2.047$, $p < 0.05$]. It has been shown that the frequency and consistency of distributional frames contributes to infants' ability to recognize adjacent words (Bortfeld et al., 2005; Mintz, 2013; Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Lepage, 2008) and is correlated with words' age of acquisition (Hills, 2013).

English also provides situations where verbs' distributional statistics become more like nouns. One such example is verbs that occur in *–ing* contexts. This situation renders a distributional context that is noun-like in its potency. If *–ing* were treated as a separate unit in language, this would make it the sixth most frequent word in the English language.[2] As a frequent anchor collocation of verbs, its co-occurrence probability with verbs is $p = 0.162$, considerably higher than the next most frequent word (*you*, $p = 0.117$) and more in line with nouns' most frequent co-occurring element (*the*, $p = 0.190$). Thus, verbs occurring in *–ing* contexts provide an ideal circumstance for testing whether distributional factors like anchor word co-occurrence are important for infant word recognition.

There are a number of other reasons to think that young infants should have an easier time recognizing verbs in *–ing* contexts, and that this difference may be enough to make the noun–verb gap disappear. Behavioral evidence suggests that older infants use highly frequent morphological units such as *–ing* as a segmentation or recognition cue. Mintz (2013) found that much older infants (18-months) are facilitated in segmenting novel sound sequences from fluent speech when they are heard in an *–ing* context (but not an unfamiliar *–dut* context). Similarly, Marquis and Shi (2008) found that French-learning 11-month-olds show facilitated recognition of French verbs in the highly frequent *–er* inflectional form (see also Marquis & Shi, 2012).

We therefore hypothesized that infants may show enhanced verb recognition, when the target verbs appear in the highly felicitous *–ing* context. If so, this would lead to three major conclusions. First, an important contribution to differences between nouns and verbs in recognition difficulty would be attributable to distributional differences of the contexts in which those words occur. Words with more useful distributional statistics would be recognizable at earlier ages. Second, although nouns *tend* to have distributional statistics that are more useful for word

[1] Statistics were calculated using the 150 most frequent nouns and 150 most frequent verbs in the CHILDES corpus (MacWhinney, 2000).

[2] As with the previous analyses, this numbers were computed using all samples of child-available speech in the CHILDES database.

Download English Version:

https://daneshyari.com/en/article/10457534

Download Persian Version:

https://daneshyari.com/article/10457534

Daneshyari.com