Cognition 138 (2015) 10–20

Contents lists available at ScienceDirect

Cognition

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

Toddlers encode similarities among novel words from meaningful sentences

Erica H. Wojcik *, Jenny R. Saffran

University of Wisconsin—Madison, Waisman Center, 1500 Highland Avenue, Madison, WI 53705, United States

ARTICLE INFO

Article history:
Received 10 April 2014
Revised 13 January 2015
Accepted 17 January 2015
Available online 19 February 2015

Keywords:
Word learning
Lexical development
Syntactic bootstrapping
Semantic networks

ABSTRACT

Toddlers can learn about the meanings of individual words from the structure and semantics of the sentences in which they are embedded. However, it remains unknown whether toddlers encode similarities among novel words based on their positions within sentences. In three experiments, two-year-olds listened to novel words embedded in familiar sentence frames. Some novel words consistently occurred in the subject position across sentences, and others in the object position across sentences. An auditory semantic task was used to test whether toddlers encoded similarities based on sentential position, for (a) pairs of novel words that occurred within the same sentence, and (b) pairs of novel words that occurred in the same position across sentences. The results suggest that while toddlers readily encoded similarity based on within-sentence occurrences, only toddlers with more advanced grammatical knowledge encoded the positional similarities of novel words across sentences. Moreover, the encoding of these cross-sentential relationships only occurred if the exposure sentences included a familiar verb. These studies suggest that the types of lexical relationships that toddlers learn depend on the child's current level of language development, as well as the structure and meaning of the sentences surrounding the novel words.

1. Introduction

How do children build a lexicon? In addition to the connections between labels and their referents, mature lexical knowledge also contains the associations between words. This lexical-semantic structure includes information gleaned from what we see and experience in the world, such as the fact that "dog" and "cat" refer to similar animals. However, there are also rich patterns in the linguistic input that young word learners could use to construct a lexicon. Words are not strung together randomly; languages are structured such that the positions of words in a sentence convey information about their meanings. Despite the fact that this structure is well documented, and that it provides potentially useful information about the semantics of words, we know very little about the types of lexical relationships that young children track in the speech stream when they hear new words.

Toddlers have an impressive ability to infer the meaning of an individual novel word from simply hearing it in a sentence. For example, toddlers can use word order and argument structure to infer the meaning of a novel verb (Naigles, 1990; Yuan & Fisher, 2009). In these studies, toddlers hear a sentence such as "*The duck is gorping the bunny*," and use the syntax of the sentence to guide their interpretation of the novel verb (measured via looking behavior; Naigles, 1990). In addition to using syntax to map a novel verb to an action, toddlers can also use syntactic structure to glean information about other semantic properties of a novel verb, such as the subjects and objects

* Corresponding author. Tel.: +1 (608) 263 5876.
E-mail address: ehwojcik@wisc.edu (E.H. Wojcik).

http://dx.doi.org/10.1016/j.cognition.2015.01.015
0010-0277/

it selects for (e.g., Gertner, Fisher, & Eisengart, 2006). Importantly, there is also evidence that this semantic information is encoded and can be used for later comprehension and learning (e.g., Yuan & Fisher, 2009). These findings demonstrate that young children readily exploit sentence structure to guide online comprehension of a novel word, and that they encode this semantic information into their lexical representations.

In addition to using sentence *structure* to learn about novel word meanings, toddlers also use the *semantics* of surrounding words to infer word meanings. For example, both adults and children use familiar verbs to predict the semantic properties of upcoming nouns (Altmann & Kamide, 1999; Fernald, Zangl, Portillo, & Marchman, 2008; Friedrich & Friederici, 2005; Valian, Prasada, & Scarpa, 2006). Moreover, children as young as 15 months of age use familiar verbs to predict semantic properties of adjacent *novel* nouns (Ferguson, Graf, & Waxman, 2014; Goodman, McDonough, & Brown, 1998; Mani & Huettig, 2012; Yuan, Fisher, Kandhadai, & Fernald, 2011).

This literature on using sentential context to infer word meaning has focused on the information that young children can use and encode about individual words. However, adult word knowledge includes much more than just the meanings of isolated words; the structure of our lexical-semantic knowledge is better described as a web or network than as a dictionary (Elman, 2009; McNamara, 2005; Rogers & McClelland, 2004). As mature language users, we are exquisitely attuned to the similarities among the meanings and functions of words in our lexicon. Representing the relationships between words, or the presence of interconnectivity or structure within the lexicon, allows us to understand and produce language flexibly and efficiently.

1.1. Gleaning word similarity from sentences

In addition to semantic information about individual words, sentential context provides information about lexical structure, or word relationships. For example, consider the following sentences: "The cat drank the milk" and "The dog drank the water." These two sentences reflect a general pattern found in English, namely that the subject of a verb phrase is often the agent of the action; "cat" and "dog" are both animate agents who are doing the drinking. Additionally, the verb *drink* selects for objects with certain semantic properties—inanimate, palatable, and liquid. Thus, if a child has sufficient knowledge of their language's grammatical structure and knows the verb *drink*, she could use this information to determine many semantic properties of the individual nouns (*cat*, *dog*, *milk*, and, *water*).

Crucially, though, a child who hears these sentences, and other sentences that use these four nouns in similar ways, could also learn about the relationships between the nouns. Over time, the child could encode that "cat" and "milk" regularly appear in the same sentence, as do "dog" and "water". This type of pattern, in which two words co-occur with an intervening item, is referred to as a non-adjacent dependency (e.g., Gómez, 2002; Newport & Aslin, 2004). Non-adjacent dependency learning is crucial to acquiring a language because many linguistic patterns follow this structure (Chomsky, 1957). Indeed, many theories purport that tracking this type of relationship is a crucial step in syntax acquisition (Gómez, 2002; Newport & Aslin, 2004).

It is possible that non-adjacent dependencies also contribute to the learning of semantic relationships. By tracking which words occur within the same sentence (albeit separated by intervening items), children could begin to learn which words occur within the same event structure. For example, "cat" and "milk" are semantically related because they participate in the same event. This type of thematic relationship between nouns is a component of mature lexical networks (e.g., Hare, Jones, Thomson, Kelly, & McRae, 2009). Thus, young word learners could use this type of within-sentence regularity to begin to build their semantic network.

In addition to within-sentence relationships, there is another type of informative structure relating the nouns in the example sentences ("The cat drank the milk" and "The dog drank the water"). Words that are used in the same position across sentences (and particularly with the same verb, e.g., cat and dog; milk and water) share semantic properties. In the current example, the subjects are animate, and the objects are inanimate potable liquids. Thus, the tracking of word positions across sentences could help children detect meaningful similarities. Indeed, adults judge words that are used in the same position across sentences as more similar to each other than words that are used in different positions (Jones & Love, 2007). Moreover, corpus analyses of child-directed speech indicate that the sentence frames surrounding words are predictive of their grammatical category (Mintz, 2003). Thus, children could potentially use these similarities in how words are used across sentences to discover the semantic relationships between words.

Within the two example sentences alone, there is complex and useful information available not only about individual nouns, but also the relationships among the four nouns. However, we do not know whether young children encode these relationships during language acquisition. This type of learning could be particularly useful when children hear novel words for the first time. Encoding the similarities in usage among novel words, as well as between novel and familiar words, could aid in the integration of those words into the developing lexicon. The current set of studies was designed to ask whether toddlers track similarities among novel words based on how they are used in sentences. Specifically, we exposed toddlers to novel words within familiar English sentence frames that provided potentially useful information about how the words' referents are related to each other. The test items were designed to determine which types of lexical relationships the toddlers encoded from those sentences.

A second aim of the current studies was to examine the relationship between word-relationship encoding and language skill. There is evidence that children with larger vocabularies and increased grammatical knowledge are better able to use distributional cues and sentence context to learn new words (e.g., Fisher, Klingler, & Song, 2006; Lany & Saffran, 2011). We were particularly interested in the connection between *grammatical knowledge* and the