Journal of Memory and Language 97 (2017) 17–29

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Distributional learning of subcategories in an artificial grammar: Category generalization and subcategory restrictions

Patricia A. Reeder [a,*], Elissa L. Newport [b], Richard N. Aslin [c]

[a] Department of Psychological Science, Gustavus Adolphus College, St. Peter, MN, USA
[b] Center for Brain Plasticity and Recovery, Georgetown University, Washington, DC, USA
[c] Department of Brain & Cognitive Sciences, University of Rochester, Rochester, NY, USA

ARTICLE INFO

Article history:
Received 27 December 2015
Revision received 12 July 2017
Available online 20 July 2017

Keywords:
Categorization
Linguistic subcategories
Language acquisition
Distributional learning

ABSTRACT

There has been significant recent interest in clarifying how learners use distributional information during language acquisition. Many researchers have suggested that distributional learning mechanisms play a major role during grammatical category acquisition, since linguistic form-classes (like *noun* and *verb*) and subclasses (like *masculine* and *feminine* grammatical gender) are primarily defined by the ways lexical items are distributed in syntactic contexts. Though recent experimental work has affirmed the importance of distributional information for category acquisition, there has been little evidence that learners can acquire linguistic *subclasses* based only on distributional cues. Across two artificial grammar-learning experiments, we demonstrate that subclasses can be acquired from distributional cues alone. These results add to a body of work demonstrating rational use of distributional information to acquire complex linguistic structures.

Introduction

Natural languages are highly structured systems, governed by particular organizational rules and representations. Language learners are tasked with acquiring these rules and representations in a primarily unsupervised environment, without initial access to the full set of sounds, word combinations, or structures that are necessary to produce and comprehend the infinite set of possible sentences in their language. One of the main linguistic structures that support a language's generativity are its syntactic categories. These form-class categories are primarily defined based on how groups of words are distributed with certain syntactic arguments. For example, certain words can occur as the subject of a *verb* or the object of a *preposition*. Words that have these syntactic properties (among others) are grouped together as *nouns*. Having the category *noun* allows a language user to use new nouns in syntactic contexts where they have previously heard other nouns occur; that is, the distributional properties of the category *noun* can be generalized across words in the category.

Languages not only have major form-class categories like *noun* and *verb*; some of these categories may be further divided into subcategories. Like major form-class categories, language subcategories are partly defined and differentiated based on the different types of linguistic contexts in which words in the subcategory may occur (e.g., Bloomfield, 1933; Chomsky, 1965; Harris, 1954). One well-studied example of noun subcategories is grammatical gender. In many languages, nouns differ in the form of the determiner that goes with them (e.g., in French, masculine nouns take the definite determiner *le*, whereas feminine nouns take the definite determiner *la*) or in the endings that must occur on the noun or on co-occurring adjectives. Importantly, linguistic gender is arbitrarily defined: grammatical gender does not clearly relate to natural biological/social gender, linguistic gender assignments are inconsistent across languages, and the number of grammatical genders in a language varies cross-linguistically. Though not all languages have grammatical gender, nouns in many languages contain other types of subcategories, such as the distinction between count nouns and mass nouns. In English, determiners serve as one type of distributional cue to these subcategories: whereas mass nouns may occur with the determiners *much* and *some*, count nouns occur with determiners such as *many* or *one*. Verbs can be subdivided based on whether or not the verb takes an object, forming transitive and intransitive subcategories, or in many languages are subdivided into conjugations, differing in the endings the verb takes for person and number. While the distinction between transitive and intransitive subcategories is related to verb semantics and argument structure, verb conjugations are distributionally defined.

[*] Corresponding author at: Department of Psychological Science, 800 West College Avenue, Gustavus Adolphus College, Saint Peter, MN 56082, USA.

E-mail addresses: preeder@gustavus.edu (P.A. Reeder), eln10@georgetown.edu (E.L. Newport), aslin@cvs.rochester.edu (R.N. Aslin).

http://dx.doi.org/10.1016/j.jml.2017.07.006
0749-596X/© 2017 Elsevier Inc. All rights reserved.

Because linguistic categories and subcategories are crucial components of natural language structure, there has been sustained interest in studying the mechanisms underlying their acquisition. However, the exact process underlying their acquisition has been particularly difficult to define. Categories and subcategories lack consistent perceptual or semantic cues to their organization, and distributional cues are often ambiguous and overlapping (e.g., Braine, 1987). Despite the complexity of this system, however, even young children demonstrate early knowledge of the form-class organization of their native language (e.g., Maratsos & Chalkley, 1980). This knowledge allows them to use syntactic categories and subcategories to learn the meanings of new words (e.g., Scott & Fisher, 2009; Yuan & Fisher, 2009) and to produce grammatical utterances based on form-class category knowledge (e.g., Berko, 1958). Even though children may not have perfect subcategory representations by the time they demonstrate productive use of form-class categories, there is evidence that they at least have basic knowledge of relevant subcategories at a very early age. For example, children acquiring the Russian gender paradigm do not consistently mark the correct gender at an early age, but they do have the correct number of gender subcategories despite occasionally using them in the wrong contexts (e.g., Gvozdev, 1961; Polinsky, 2008). Thus, although there may be imperfect production of subcategory knowledge (perhaps due to performance limitations), grammatical subcategories are clearly being formed early in language development (e.g., Valian, 1986).

Given the potential complexity of category acquisition, a large body of work has explored the types of information that learners could in principle – and do in practice – exploit for discovering the categories and subcategories in their language. Though natural language categories are associated with many possible sources of cues, distributional information has proven to be a reliable cue to major form class category structure (e.g., Cartwright & Brent, 1997; Mintz, 2003; Mintz, Newport, & Bever, 2002; Redington, Chater, & Finch, 1998). Additionally, human learners have been shown to use the distributional cues that define categories – sometimes along with other types of cues – in order to acquire them (e.g., Braine et al., 1990; Brooks, Braine, Catalano, Brody, & Sudhalter, 1993; Mintz, 2002; Mintz, Wang, & Li, 2014; Reeder, Newport, & Aslin, 2013; Schuler, Reeder, Newport, & Aslin, in press; Scott & Fisher, 2009; St. Clair, Monaghan, & Christiansen, 2010).

A first step in studying the role of distributional information for categorization was provided by Smith (1966), who showed that learners were quite capable of learning a simple language consisting of two categories:

Pair → $\alpha + \beta$
α → D, V, H, R, X
β → M, F, G, K, L

where there are two categories of letters (α and β) and one rule that requires α words to be followed by β words. Participants saw some of the possible strings of the language and were then asked to recall as many strings as possible. The results showed that learners recalled both the presented strings and ‘‘intrusions” (legal strings according to the pairing rule of the language that were not presented during exposure). The recall of grammatical intrusions is evidence of category-level generalizations, where the categories are defined by positional information (the co-occurrence statistics between the two categories were distributionally uninformative in this study).

However, in a similar paradigm by Smith (1969), participants had to learn dependencies between words within a pair of contingent categories:

Pair → $\alpha + \beta$
α → M, P
β → N, Q
M → m_1, m_2, m_3
N → n_1, n_2, n_3
P → p_1, p_2, p_3
Q → q_1, q_2, q_3

Importantly, strings of the language followed the basic pattern of MN or PQ; no MQ or PN strings were presented. M, N, P, and Q were categories of 3 items (letters) each. Exposure consisted of seeing 2/3rds of the possible MN pairings and 2/3rds of the PQ pairings. However, while participants learned that M- and P-words occurred first and that N- and Q-words occurred last in the 2-word strings of the language, they did not learn the co-occurrence dependencies that M-words were only followed by N-words and P-words were only followed by Q-words. They produced MQ strings as well as PQ strings, and showed no differentiation between the two. This ‘‘MN/PQ problem” (Braine, 1987) is a classic case, widely cited in the literature, of failure to acquire categories from distributional information alone.

Other problems have also plagued learning theories that primarily rely on distributional analyses for category formation. As Pinker (1984, 1987) noted, it is not always obvious which contexts a learner should learn from in any particular utterance, and overly simplistic distributional analyses could lead a learner astray. Likewise, Braine (1987) recognized how easily and quickly learners acquired positional cues to categories in the MN/PQ problem, such as ‘‘M-words come first” and ‘‘N-words come last.” Though positional cues are a type of distributional information, they do not reveal the full set of rules governing the MN/PQ language. Unfortunately for proponents of distributional analyses, it seemed as if learners were *only* capable of acquiring these serial dependencies in Smith (1969), since they were unable to learn the rule ‘‘M words are obligatorily followed by N-words.” Braine (1987) concluded that learners required an additional salient cue (called a ‘‘similarity relation”) to overcome this positional information and highlight the distributional structure of the categories in the MN/PQ problem – for example, associating the M subclass with males and the P subclass with females, thus building in a semantic similarity relation. With the addition of partially correlated semantic cues, subjects were able to restrict generalization in the MN/PQ experiment: they made fewer ungrammatical overgeneralizations when a semantic similarity relation cued them into the co-occurrence structure of the MN/PQ subclasses.

A number of investigators have followed up on this hypothesis, exploring the role of shared cues to category structure (e.g., Braine, 1966; semantic cues: Braine et al., 1990; morphological cues: Brooks et al., 1993; phonological cues: Frigo & McDonald, 1998; Gerken, Gomez, & Nurmsoo, 1999; Gerken, Wilson, & Lewis, 2005; Monaghan, Chater, & Christiansen, 2005; Morgan, Shi, & Allopenna, 1996; Shi, Morgan, & Allopenna, 1998; Wilson, 2002; shared features: Gomez & Lakusta, 2004). The results from many of these artificial language studies suggest that the formation of linguistic classes crucially depends on overlapping perceptual properties that link the items together. These correlated perceptual cues might arise from identity or repetition of elements in grammatical sequences, or from a phonological or semantic cue identifying words across different sentences as similar to one another (for example, words ending in *–a* are feminine). On this view, correlated cues are necessary and sufficient to discover the categorical structure in artificial languages, and in the acquisition of natural grammatical classes (Gomez & Gerken, 2000).

However, most categories (and most subcategories) are arbitrary: though they may have partially correlated semantic,