Review

Word classes in the brain: Implications of linguistic typology for cognitive neuroscience

David Kemmerer [a,b,*]

[a] Department of Speech, Language, and Hearing Sciences, Purdue University, United States
[b] Department of Psychological Sciences, Purdue University, United States

ARTICLE INFO

ABSTRACT

Although recent research on the neural substrates of word classes has generated some valuable findings, significant progress has been hindered by insufficient attention to theoretical issues involving the nature of the lexical phenomena under investigation. This paper shows how insights from linguistic typology can provide cognitive neuroscientists with well-motivated guidelines for interpreting the extant data and charting a future course. At the outset, a fundamental distinction is made between universal and language-particular aspects of word classes. Regarding universals, prototypical nouns involve reference to objects, and their meanings rely primarily on the ventral temporal lobes, which represent the shape features of entities; in contrast, prototypical verbs involve predication of actions, and their meanings rely primarily on posterior middle temporal regions and frontoparietal regions, which represent the visual motion features and somatomotor features of events. Some researchers maintain that focusing on object nouns and action verbs is inappropriate because it conflates the semantic and grammatical properties of each word class. However, this criticism not only ignores the importance of the universal prototypes, but also mistakenly assumes that there are straightforward morphological and/or syntactic criteria for identifying nouns and verbs in particular languages. In fact, at the level of individual languages, the classic method of distributional analysis leads to a proliferation of constructionally based entity-denoting and event-denoting word classes with mismatching memberships, and all of this variation must be taken seriously, not only by linguists, but also by cognitive neuroscientists. Many of these word classes involve remarkably close correspondences between grammar and meaning and hence are highly relevant to the neurobiology of conceptual knowledge, but so far hardly any of them have been investigated from a neurolinguistic perspective.

* Department of Speech, Language, and Hearing Sciences, 500 Oval Drive, Purdue University, West Lafayette, IN 47907-2038, United States.
E-mail address: kemmerer@purdue.edu.

1. Introduction

It has been recognized for centuries that word classes—also known as form classes, lexical categories, syntactic categories, grammatical categories, and parts of speech—are fundamental to the grammatical systems of human languages. The most familiar word classes have traditionally been called nouns, verbs, and adjectives, but many others have also been posited, such as prepositions, conjunctions, complementizers, determiners, and so on. Because these sorts of word classes play central roles in the morphological and syntactic phenomena of languages worldwide, they are explicitly addressed, in one way or another, in all linguistic theories, and they figure prominently in many of the extended branches of the language sciences that deal with the representation and processing of lexical knowledge, including computational linguistics, psycholinguistics, and, of special interest here, neurolinguistics.

A substantial amount of neurolinguistic research has focused on how word classes are implemented in the brain. Most of this work has concentrated on what are widely regarded as the two major word classes, namely nouns and verbs, and in fact this basic distinction has been investigated with all of the main brain-mapping methods: deficit-lesion correlations in brain-damaged patients; positron emission tomography (PET); functional magnetic resonance imaging (fMRI); extracranial and intracranial electrophysiology; magnetoencephalography (MEG); and transcranial magnetic stimulation (TMS) (for reviews and meta-analyses see Berlingeri et al., 2008; Black & Chiat, 2003; Cappa & Perani, 2003; Crepaldi et al., 2013; Crepaldi, Berlingeri, Paulesu, & Luzzatti, 2011; Druks, 2002; Luzzatti, Aggujaro, & Crepaldi, 2006; Mätzig, Druks, Masterson, & Vigliocco, 2009; Pillon & d'Honincthun, 2010; Shapiro & Caramazza, 2003b, 2004, 2009; Vigliocco, Vinson, Druks, Barber, & Cappa, 2011). The literature on this topic is now quite large, and it contains a wealth of intriguing ideas and discoveries with broad ramifications. For example, two of the most influential researchers, Shapiro and Caramazza (2003b, 2004, 2009), have argued that the noun–verb distinction is reflected in multiple isolable aspects of the neural organization of language: the meanings of words (e.g., object concepts vs action concepts); the morphological and syntactic operations that apply to words (e.g., inflection of nouns for number vs inflection of verbs for tense); and the pathways that connect those semantic and grammatical systems with the phonological and orthographic systems (as explored by studies of brain-damaged patients who have production deficits that are not only grammatical category-specific, being restricted to either nouns or verbs, but also modality-specific, being restricted to either spoken or written output).

But even though this field of research has made significant progress in recent years, it has also generated many confusing, seemingly contradictory sets of results, and it has failed to resolve a number of long-standing controversies, especially regarding the precise nature of the relationship between the semantic and grammatical properties of word classes. In this paper, I will argue that these problems are due largely to insufficient consideration of the actual definitions of the technical terms that are typically used to refer to the word classes under investigation—that is, terms like noun and verb. In the majority, perhaps even the vast majority, of neurolinguistic studies on word classes, the authors do not provide detailed, theoretically grounded definitions of these terms; instead, they generally assume, often implicitly, that the meanings are straightforward, or at least clear enough for the purpose of the given experiment. While there are certainly some exceptions to this trend [e.g., see the "general conclusions" of Vigliocco et al.'s (2011) review], the tendency to be vague is strong, and the unfortunate effect is that the central scientific notions in the field are rarely unpacked; instead, they are just transferred from one study to the next, like the unexamined contents of a locked suitcase being passed from one person to another.

This state of affairs is worrisome, because if the field as a whole does not open up that suitcase, so to speak, and adopt a more elaborate, coherent, and well-justified framework for characterizing word classes at a strictly linguistic level of analysis, there is a good chance that it will continue to be hindered in various ways. For one thing, the interpretation of much of the extant data may remain just as theoretically under-constrained and open to debate as it has been up to now. In addition, a large proportion of future studies may persist in searching for the neural correlates of word classes that are only specified rather amorphously without reference to relevant research in linguistics. On the other hand, if the field as a whole does make a more concerted effort to increase the sophistication of its theoretical treatment of word classes, it will likely benefit a great deal. For instance, current controversies may become more amenable to resolution from a fresh perspective, and future studies may be designed with the aim of revealing the neural correlates of word classes whose unique semantic and grammatical properties have already been precisely specified by careful linguistic research.

Of course, this raises the important question of which framework for characterizing word classes is most suitable for adoption in neurolinguistics. Not surprisingly, the answer is by no means obvious, since word classes are inherently complicated phenomena, and many competing accounts have been proposed by contemporary generative, cognitive, functional, and typological theories (for a survey see Rauh, 2010). All of these frameworks undoubtedly have much to offer neurolinguistics, but in this paper I will focus primarily on typological approaches—that is, approaches that have the distinct advantage of being anchored in the investigation of similarities and differences among the roughly 6000 languages of the world. Now, even in the relatively narrow domain of typological research on word classes, several alternative frameworks have been proposed (Bisang, 2011; Dixon, 2010b; Haspelmath, 2007, 2012b; Hengeveld, 1992; Hengeveld & van Lier, 2010; Rijkhoff & van Lier, 2013; Schacter, 1985; Schacter & Shopen, 2007; Van Valin, 2008; Vogel & Comrie, 2000). Here I will concentrate on the theory developed by Croft (1991, 2000a, 2000b, 2001, 2005, 2007a, 2007b, 2009, 2010a, 2010b, 2012, 2013, in press; Croft & van Lier, 2012), while drawing on other work when appropriate. Croft's theory is especially useful for present purposes not only because it has a long and respectable history in the field, but also because it takes into account both morphosyntactic