

Oxford  
LINGUISTICS

# Linguistic Categorization

John R. Taylor

THIRD EDITION

# catego

OXFORD TEXTBOOKS IN LINGUISTICS

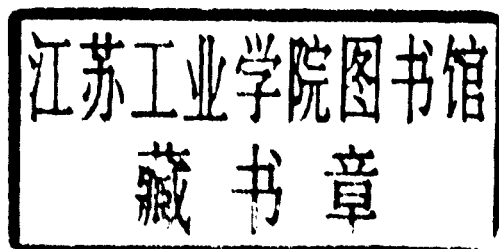
---

# Linguistic Categorization

---

*Third edition*

John R. Taylor



OXFORD  
UNIVERSITY PRESS

# OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.

It furthers the University's objective of excellence in research, scholarship,  
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai

Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata

Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi

São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press  
in the UK and in certain other countries

Published in the United States

by Oxford University Press Inc., New York

© John R. Taylor 1989, 1995, 2003

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2003

All rights reserved. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means,  
without the prior permission in writing of Oxford University Press,  
or as expressly permitted by law, or under terms agreed with the appropriate  
reprographics rights organization. Enquiries concerning reproduction  
outside the scope of the above should be sent to the Rights Department.  
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover  
and you must impose this same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Data available

ISBN 0-19-926664-6

10 9 8 7 6 5 4 3 2 1

Typeset in Times and Meta

by RefineCatch Limited, Bungay, Suffolk

Printed in Great Britain by

T.J. International Ltd, Padstow, Cornwall

OXFORD TEXTBOOKS IN LINGUISTICS

*Series editors*

Keith Brown, Eve V. Clark, April McMahon,  
Jim Miller, Lesley Milroy

---

# Linguistic Categorization

---

## OXFORD TEXTBOOKS IN LINGUISTICS

*General editors:*

**Keith Brown**, University of Cambridge; **Eve V. Clark**, Stanford University;  
**April McMahon**, University of Sheffield; **Jim Miller**, University of Edinburgh;  
**Lesley Milroy**, University of Michigan

This series provides accessible and authoritative textbooks on the approaches, methods, and theories associated with the main subfields of linguistics.

### PUBLISHED

**A Practical Introduction to Phonetics** (SECOND EDITION)

*by J. C. Catford*

**Meaning in Language:** An Introduction to Semantics and Pragmatics

*by Alan Cruse*

**Principles and Parameters:** An Introduction to Syntactic Theory

*by Peter W. Culicover*

**Linguistic Reconstruction:** An Introduction to Theory and Method

*by Anthony Fox*

**Semantic Analysis:** A Practical Introduction

*by Cliff Goddard*

**Cognitive Grammar**

*by John R. Taylor*

### IN PREPARATION

**The Grammar of Words:** An Introduction to Linguistic Morphology

*by Professor G. Booij*

**Pragmatics**

*by Yan Huang*

# Preface to the third edition

The third edition of *Linguistic Categorization* represents a thorough revision of the earlier editions. A number of errors, obscurities, and stylistic infelicities of earlier editions have been removed; new material has been added throughout; a new chapter (Chapter 8) has been added; the contents of Chapter 14 of the second edition have been repositioned elsewhere; bibliographical references have been updated; and each chapter has been supplemented by study questions and suggestions for further reading.

J. R. T.

# Typographical conventions

Linguistic forms are printed in italics: *dog*

Meanings of linguistic forms, and glosses of foreign language forms, are given between double quotes: “dog”

Citations are marked by single quotes.

Names of categories are printed in small capitals: DOG

Phonetic and semantic features are printed in small capitals enclosed in square brackets: [VOCALIC], [ANIMAL]

Semantic attributes are printed in normal type enclosed in square brackets: [ability to fly]

Phonemes, and phonemic transcriptions, are enclosed in slashes //

Phonetic symbols and phonetic transcriptions are enclosed in square brackets [ ]

An asterisk \* indicates that a following linguistic expression is unacceptable, on either semantic or syntactic grounds. Expressions of questionable acceptability are preceded by a question mark.

**Bold-face** is used for the first appearance of technical terms.

# Introduction and Overview

All living creatures, even the lowliest, possess the ability to categorize. In order to survive, a creature has to be able, at the very least, to distinguish what is edible from what is inedible, what is benign from what is harmful. And in order to mate and reproduce, a creature must be able to recognize its own kind.

Strictly speaking, every entity and every situation that we encounter is uniquely different from every other. In order to function in the world, all creatures, including humans, need to be able to group different entities together as instances of the same kind. Our cognitive apparatus does this for us automatically, most of the time. We ‘automatically’ classify things around us as ‘books’, ‘pencils’, ‘trees’, ‘coffee-cups’, and so on. And having classified something as a ‘book’, ‘coffee-cup’, or whatever, we have access to further knowledge about the object—we know how to handle it, what we are supposed to do with it, and so on. In this way, categorization serves to reduce the complexity of the environment. For the most part, categorization seems part and parcel of perception. Seeing something, and categorizing it as, say, a pencil, seem one and the same process. And if we come across something whose categorization is not immediately evident, the experience tends to create a feeling of unease. ‘Well, what *is* it?’, we ask.

Human beings are categorizing creatures par excellence. Our ability to function in the complex physical and social world in which we find ourselves depends on elaborate categorizations of things, processes, persons, institutions, and social relations. We are able to create and operate with literally tens of thousands, perhaps hundreds of thousands of categories, ranging from the extremely fine-grained to the highly general. Moreover, categorization is flexible, in that we can modify existing categories in order to accommodate new experiences, and we can create new categories whenever the need arises.

Linguists are interested in categorization for two reasons:

(i) Categorizing something very often involves naming it. In fact, the meaning of a word can often be taken to be the name of a category. To know the word *tree* is to know, amongst other things, what counts as a tree. In virtue of knowing the word, we are able to apply it correctly to anything that can



be categorized as a tree (in contrast, say, to a bush, a shrub, or an electricity pylon). To a large extent, the study of lexical semantics *is* the study of categorization. But it is not only the words of a language which denote categories. Grammatical phenomena, such as the tense of a verb, or the type of clause in which a verb occurs, may also be associated with meanings, and these meanings, likewise, can be thought of as categories. To use a verb in the past tense, or to describe a situation using a transitive clause, brings with it a certain categorization of the described state of affairs.

(ii) Second, language itself is an object of categorization. Like everything else in our environment, every linguistic utterance is a unique event. That language is able to be used as a means of communication at all rests on our ability to recognize, in the flux of language events, instances of linguistic categories—such things as phonemes, words, word classes, clause types, and so on. Thus, acoustically different sounds get categorized as instances of the same phoneme, sound-meaning relations get categorized as instances of the same word, diverse linguistic expressions get categorized as examples of the same lexical or syntactic category, such as NOUN, VERB, TRANSITIVE CLAUSE.

The title of this book, *Linguistic Categorization*, is meant to reflect this double role of categorization in the study of language. In one of its senses, the title refers to the process by which people, in using language, are categorizing their experience of the world. Whenever we use the word *dog* to refer to different animals, or describe different colour sensations by the same word, such as *red*, we are undertaking acts of categorization. Although individually different, the animals and the colours are regarded as being examples of the ‘same’ categories. Much of this book, therefore, is about the meanings of linguistic forms, and the categorization of the world which a knowledge of these meanings entails.

The title of the book is also to be understood to refer to the categories of language itself. Chapters 11 to 13 focus in particular on the categorization of linguistic objects. We will find that the categories of linguistic objects are structured along much the same lines as non-linguistic objects. Insights gained from the study of semantic categories such as RED and BIRD can be profitably applied to the study of linguistic categories such as PHONEME, WORD, and TRANSITIVE CLAUSE.

## Where do categories come from?

What is the basis on which two distinct entities come to be regarded as instances of the same category? We can distinguish two, radically opposed approaches.

One extreme position, that of **nominalism**, claims that sameness is merely a matter of linguistic convention; the range of entities which may be called dogs, or the set of colours that may be described as red, have in reality nothing in common but their name. We call them by the same name, simply because that is the linguistic convention. An equally extreme position is that of **realism**. Realism claims that categories like *DOG* and *RED* exist independently of language and its users, and that the words *dog* and *red* merely name these pre-existing categories. An alternative position is **conceptualism**. Conceptualism postulates that a word and the range of entities to which it may refer are mediated by a mental entity, i.e. a concept. It is in virtue of our knowledge of the concepts “dog” and “red”, i.e. in virtue of our knowledge of the meanings of the words *dog* and *red*, that we are able to categorize different entities as dogs, different colours as red, and so on. Conceptualism may be given a nominalist or a realist orientation. On the one hand, we might assume that concepts simply reflect prevailing linguistic conventions. On this view, the English speaker’s concepts “red” and “dog” arise through observation of how the words *red* and *dog* are conventionally used; once formed, the concepts will govern future linguistic performance. Alternatively, we might claim that concepts mirror really existing properties of the world. On this view, our concepts are not arbitrary creations of language, but constitute part of our understanding of what the world is ‘really’ like.

This book will take a course which is intermediate between these two positions, yet strictly speaking consonant with neither. To the extent that a language is a set of conventionalized pairings of forms and meanings, it may indeed be the case that a language imposes a set of categories on its users. That something is conventionalized, however, does not entail that it is arbitrary. The categories encoded in a language may be motivated, to varying degrees, by a number of factors—by actually existing discontinuities in the world, by the manner in which human beings interact, in a given culture, with the world, and by general cognitive processes of concept formation. It is precisely the dialectic of convention and motivation which gives rise to the fact that the categories encoded in one language do not always stand in a one-to-one correspondence with the categories of another language. Languages are indeed diverse with respect to the categories which they encode; yet the diversity is not unconstrained.

## Background

The theoretical background to this book is a set of principles and assumptions that have come to be known as ‘cognitive linguistics’. Cognitive linguistics first emerged as a coherent theoretical approach in the mid 1980s, largely as a reaction against the excesses of ‘formalist’ linguistics in the Chomskyan mould. One of the main points of contention concerns the nature of linguistic

knowledge. Whereas Chomskyan linguists regard knowledge of language as an autonomous component of the mind, independent, in its essential constitution, from other kinds of knowledge and from other cognitive abilities, cognitive linguists posit an intimate, dialectic relationship between the structure and function of language on the one hand, and non-linguistic skills and knowledge on the other. Language, being at once both the creation of human cognition and an instrument in its service, is thus more likely than not to reflect, in its structure and functioning, more general cognitive abilities. One of the most important of these cognitive abilities is precisely the ability to categorize, i.e. to see sameness in diversity. A study of categorization processes is thus likely to provide valuable insights into the meanings symbolized by linguistic forms. Furthermore, there is every reason to expect that the structural categories of language itself will be analogous, in many ways, to the categories which human beings perceive in the nonlinguistic world around them.

Specifically, this book explores the implications for linguistic categorization (in both of the senses of the term) of the notion of prototype categories. The basic idea—supported by ample empirical evidence—is that categories are structured around ‘prototypical instances’, or ‘best examples’ of the category. It follows that categories have an internal structure—some members are better examples than others. Moreover, it is often the case that categories have no clear boundaries. There may be borderline cases, where clear, unambiguous categorization is not possible. Thus, an entity may be a marginal example of more than one category, but a good example of none.

## Overview

This book is structured as follows. Chapter 1 introduces some basic issues for the study of categorization on the example of colour categories. Chapter 2 surveys the ‘classical’ model of categories, according to which categories are defined in terms of a set of necessary and sufficient conditions. Some problematic aspects of this approach are reviewed. Chapters 3 and 4 introduce the alternative prototype view. Chapter 5 emphasizes the grounding of categories in our knowledge and beliefs about the world, and promotes an essentially encyclopaedic view of linguistic meaning.

Chapter 6 broaches the topic of polysemy—the fact that most words in a language have more than one semantic value. The various meanings of a word constitute a special kind of category, namely, a family resemblance, or radial category. Chapter 7 addresses two of the principal means whereby words can acquire additional meanings, namely metonymy and metaphor. Chapter 8 returns to the question of polysemy, and asks whether it is indeed possible in all cases to neatly enumerate the different meanings of a word. Chapter 9 extends the study beyond the meanings of words and examines the meanings

of more ‘grammatical’ categories, such as the diminutive and the past tense, while Chapter 10 addresses the semantic aspects of intonation categories.

Chapters 11–13 turn to the categories of language itself—words and word classes (Chapter 11), syntactic constructions (Chapter 12), and phonological categories (Chapter 13). The topic of the final chapter is acquisition, where particular attention is paid to the acquisition of some of the categories discussed in earlier chapters of the book.

# Contents

<i>Preface to the third edition</i>	viii
<i>Typographical conventions</i>	ix
<i>Introduction and Overview</i>	xi
<b>1 The Categorization of Colour</b>	<b>1</b>
<b>1.1</b> Why colour terms?	2
<b>1.2</b> Arbitrariness	5
<b>1.3</b> An alternative approach: focal colours	8
<b>1.4</b> Autonomous linguistics vs. cognitive linguistics	14
<b>2 The Classical Approach to Categorization</b>	<b>19</b>
<b>2.1</b> Aristotle	20
<b>2.2</b> The classical approach in linguistics: phonology	22
<b>2.3</b> The classical approach in semantics	27
<b>2.4</b> What's wrong with the classical theory?	35
<b>3 Prototype Categories: I</b>	<b>41</b>
<b>3.1</b> Wittgenstein	42
<b>3.2</b> Prototypes: an alternative to the classical theory	43
<b>3.3</b> Basic level terms	48
<b>3.4</b> Prototypes and the basic level	53
<b>3.5</b> Where do prototypes come from?	55
<b>3.6</b> Some applications	58
<b>4 Prototype Categories: II</b>	<b>63</b>
<b>4.1</b> Prototypes	63
<b>4.2</b> Prototypes and schemas	69
<b>4.3</b> Folk categories and expert categories	72
<b>4.4</b> Hedges	77

<b>5</b>	<b>Linguistic and Encyclopaedic Knowledge</b>	<b>84</b>
5.1	The dictionary and the encyclopaedia	85
5.2	Domains and schemas	87
5.3	Frames and scripts	90
5.4	Perspectivization	93
5.5	Frames and scripts in the comprehension of nominals	94
5.6	Fake	96
5.7	Real	98
<b>6</b>	<b>Polysemy and Meaning Chains</b>	<b>102</b>
6.1	Monosemous and polysemous categories	102
6.2	An illustration: <i>Climb</i>	108
6.3	Over	112
6.4	Some problems	118
<b>7</b>	<b>Category Extension: Metonymy and Metaphor</b>	<b>124</b>
7.1	Metonymy	124
7.2	Metaphor	132
<b>8</b>	<b>Polysemy, or: How Many Meanings Does a Word <i>Really</i> Have?</b>	<b>144</b>
8.1	Polysemy and compositionality	145
8.2	The two-level approach	148
8.3	Two illustrations: <i>in</i> and <i>round</i>	151
8.4	Polysemy and the network model	159
<b>9</b>	<b>Polysemous Categories in Morphology and Syntax</b>	<b>170</b>
9.1	Case	170
9.2	The diminutive	172
9.3	The past tense	176
9.4	Yes–no questions	181
<b>10</b>	<b>Polysemous Categories in Intonation</b>	<b>186</b>
10.1	The problem of intonational meaning	186
10.2	The meanings of falling and rising tones	188
10.3	High key	195
<b>11</b>	<b>Grammatical Categories</b>	<b>200</b>
11.1	Words, affixes, and clitics	202
11.2	Grammatical categories	208
11.3	The semantic basis of grammatical categories	216
<b>12</b>	<b>Syntactic Constructions as Prototype Categories</b>	<b>222</b>
12.1	Why we need constructions	223
12.2	Elements of a construction grammar	225

<b>12.3</b>	The prenominal possessive	228
<b>12.4</b>	The transitive construction	231
<b>12.5</b>	The transitive construction: more marginal members	235
<b>12.6</b>	Metaphorical extension of syntactic constructions	239
<b>12.7</b>	A comparison with German	241
<b>12.8</b>	Concluding remarks	243
<b>13</b>	Prototype Categories in Phonology	247
<b>13.1</b>	Phoneme categories	248
<b>13.2</b>	The gradience of phonetic features	255
<b>13.3</b>	Syllable constructions	259
<b>14</b>	The Acquisition of Categories	266
<b>14.1</b>	Hypothesized acquisition routes	267
<b>14.2</b>	Grammatical categories and constructions	270
<b>14.3</b>	Conceptual development	275
<b>14.4</b>	Word meanings	279
	<i>References</i>	285
	<i>Author index</i>	301
	<i>Subject index</i>	305

## CHAPTER 1

# The Categorization of Colour

<b>1.1 Why colour terms?</b>	2	<b>1.4 Autonomous linguistics vs. cognitive linguistics</b>	14
<b>1.2 Arbitrariness</b>	5	<b>Study questions</b>	18
<b>1.3 An alternative approach: focal colours</b>	8	<b>Further reading</b>	18

As pointed out in the Introduction, linguistics is concerned with categorization on two levels. On the one hand, linguists need categories in order to describe the object of investigation. In this, linguists proceed just like practitioners of any other discipline. The noises that people make are categorized as linguistic or non-linguistic noises; linguistic noises are categorized as instances of a particular language, or of a dialect of a particular language; sentences are categorized as grammatical or ungrammatical; words are categorized as nouns or verbs; sequences of words are categorized as examples of syntactic constructions; sound segments are classified as vowels or consonants, stops or fricatives, and so on.

But linguists need to be concerned with categorization at another level. Many of the things that linguists study—words, morphemes, syntactic structures, intonation patterns—not only constitute categories in themselves, they also stand for categories. The phonetic form [ɪed] can not only be categorized as, variously, an English word, an adjective, a syllable with a consonant-vowel-consonant structure; [ɪed] is used by English speakers to designate a range of physically and perceptually distinct properties of the real world (more precisely, a range of distinct colour sensations caused by the real-world properties) and assigns this range of colours to the category RED. Likewise, the morphosyntactic category PAST TENSE (usually) categorizes states of affairs with respect to their anteriority to the moment of speaking; the preposition



*on* (in some of its uses) categorizes the relationship between entities as one of contact and support, and so on.

Both in its methodology and in its substance, then, linguistics is intimately concerned with categorization. The point has been made by Labov (1973: 342): 'If linguistics can be said to be any one thing it is the study of categories: that is, the study of how language translates meaning into sound through the categorization of reality into discrete units and sets of units.' Questions like: Do categories have any basis in the real world or are they merely constructs of the human mind? What is their internal structure? How are categories learnt? How do people go about assigning entities to a category? What kinds of relationships exist amongst categories? must inevitably be of vital importance to linguists. Labov, in the passage just referred to, goes on to point out that categorization 'is such a fundamental and obvious part of linguistic activity that the properties of categories are normally assumed rather than studied'. In recent years, however, research in the cognitive sciences, especially cognitive psychology, has forced linguists to make explicit, and in some cases to rethink, their assumptions. In this first chapter, I will introduce some of the issues involved, taking as my cue the linguistic categorization of colour.

## 1.1 Why colour terms?

There are good reasons for starting with colour terms. In many respects colour terminology provides an ideal testing ground for theories of categorization. It is sometimes asserted—by linguists, anthropologists, and others—that categories have neither a real-world nor a perceptual base. Reality is merely a diffuse continuum, and our categorization of it is ultimately a matter of convention, i.e. of learning. Benjamin Lee Whorf stated this view in the following well-known and often-cited passage:

The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds—and this means largely by the linguistic systems in our minds. We cut up nature, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way—an agreement that holds throughout our speech community and is codified in the patterns of our language. (Whorf 1956 [1940], 213–14)

This view was echoed by the anthropologist, Edmund Leach:

I postulate that the physical and social environment of a young child is perceived as a continuum. It does not contain any intrinsically separate 'things'. The child, in due course, is taught to impose upon this environment a kind of discriminating grid which serves to distinguish the world as being composed of a large number of separate things, each labelled with a name. This world is a representation of our