

9461519

THE LEXICON IN ACQUISITION

EVE V. CLARK

Stanford University



E9461519



CAMBRIDGE
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1993

First published 1993

Printed in Great Britain at the University Press, Cambridge

A catalogue record for this book is available from the British Library

Library of Congress cataloguing in publication data

Clark, Eve V.

The lexicon in acquisition / Eve V. Clark.

p. cm. – (Cambridge studies in linguistics: 65)

Includes bibliographical references and index.

ISBN 0 521 44050 5

1. Language acquisition. 2. Vocabulary. I. Title. II. Series.

P118.C548 1993

401'.93–dc20 92–40122 CIP

ISBN 0 521 440505 hardback

Words are the starting point. Without words, children can't talk about people, places, or things, about actions, relations, or states. Without words, children have no grammatical rules. Without words, there would be no sound structure, no word structure, and no syntax. The lexicon, then, is central in language, and in language acquisition.

Eve Clark's book argues for this centrality and for the general principles of conventionality and contrast at the core of language acquisition. She looks at the hypotheses children draw on about possible word meanings, and how they map their meanings onto forms. She starts with children's emerging knowledge of conventional words and their meanings – the ontological categories they rely on for early meanings and their strategies for mapping meanings onto forms. She then takes up their growing knowledge of word structure as reflected in their formation of new words, and shows that children learning different languages follow similar paths as they learn about words and word structure.

The lexicon in acquisition is unusual in dealing with data from a large variety of languages, in its emphasis on the general principles children rely on as they analyse complex word-forms (transparency of meaning, simplicity of form, and productivity), and in the broad perspective it takes on lexical acquisition.

CAMBRIDGE STUDIES IN LINGUISTICS

General Editors: J. BRESNAN, B. COMRIE, W. DRESSLER,
R. HUDDLESTON, R. LASS, D. LIGHTFOOT, J. LYONS,
P. H. MATTHEWS, R. POSNER, S. ROMAINÉ, N. V. SMITH,
N. VINCENT

The lexicon in acquisition

In this series

-
- 35 MARTIN ATKINSON *Explanations in the study of child language development*
 - 36 SUZANNE FLEISCHMAN *The future in thought and language*
 - 37 JENNY CHESHIRE *Variation in an English dialect*
 - 38 WILLIAM A. FOLEY and ROBERT VAN VALIN JR *Functional syntax and Universal Grammar*
 - 39 MICHAEL A. COVINGTON *Syntactic theory in the High Middle Ages*
 - 40 KENNETH J. SAFIR *Syntactic chains*
 - 41 J. MILLER *Semantics and syntax*
 - 42 H. C. BUNT *Mass terms and model-theoretic semantics*
 - 43 HEINZ J. GIEGERICH *Metrical phonology and phonological structure*
 - 44 JOHN HAIMAN *Natural syntax*
 - 45 BARBARA M. HORVATH *Variation in Australian English: the sociolects of Sydney*
 - 46 GRANT GOODALL *Parallel structures in syntax: coordination, causatives, and restructuring*
 - 47 JOHN M. ANDERSON and COLIN J. EWEN *Principles of dependency phonology*
 - 48 BARBARA A. FOX *Discourse structure and anaphora*
 - 49 LAUREL J. BRINTON *The development of English aspectual systems*
 - 50 DONNA JO NAPOLI *Predication theory*
 - 51 NOEL BURTON-ROBERTS *The limits to debate: a revised theory of semantic proposition*
 - 52 MICHAEL S. ROCHEMONT and PETER W. CULICOVER *English focus constructions and the theory of grammar*
 - 53 PHILIP CARR *Linguistic realities: an automatist metatheory for the generative enterprise*
 - 54 EVE E. SWEETSER *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*
 - 55 REGINA BLASS *Relevance relations in discourse: a study with special reference to Sissala*
 - 56 ANDREW CHESTERMAN *On definiteness: a study with special reference to English and Finnish*
 - 57 ALESSANDRA GIORGI and GIUSEPPE LONGOBARDI *The syntax of noun phrases: configuration, parameters and empty categories*
 - 58 MONIK CHARETTE *Conditions on phonological government*
 - 59 M. H. KLAIMAN *Grammatical voice*
 - 60 SARAH M. B. FAGAN *The syntax and semantics of middle constructions: a study with special reference to German*
 - 61 ANJUM P. SALEEMI *Universal Grammar and language learnability*
 - 62 STEPHEN R. ANDERSON *A-Morphous Morphology*
 - 63 LESLEY STIRLING *Switch reference and discourse representation*
 - 64 HENK J. VERKUYL *A theory of aspectuality: the interaction between temporal and atemporal structure*
 - 65 EVE V. CLARK *The lexicon in acquisition*

Supplementary volumes

- TERENCE MCKAY *Infinitival complement clauses in German*
- STELLA MARIS BORTONI-RICARDO *The urbanization of rural dialect speakers: a socio-linguistic study in Brazil*
- RUDOLF P. BOTHA *Form and meaning in word formation: a study of Afrikaans reduplication*
- AYHAN AKSU-KOÇ *The acquisition of aspect and modality: the case of past reference in Turkish*
- MICHAEL O. SIADHAIL *Modern Irish: grammatical structure and dialectal variation*
- ANNICK DE HOUWER *The acquisition of two languages from birth: a case study*
- LILIANE HAEGEMAN *Theory and description in generative syntax. A case study in West Flemish*

Earlier issues not listed are also available

Acknowledgments

This book has changed enormously since I began writing. The final version owes more than I can say to colleagues, friends, and students. I have benefited immeasurably from working with Ruth A. Berman, Kathie L. Carpenter, Sophia R. Cohen, Susan A. Gelman, Barbara F. Hecht, and Randa C. Mulford. Colleagues and friends have generously added to my collections of children's coinages over the years; and I am particularly grateful to Melissa Bowerman, Fortunée Kayra-Stuart, and Randa C. Mulford who have been more than generous in making data from their own records available to me. I also owe thanks to Elaine S. Andersen, Jean Berko Gleason, William Brewer, Michel Denis, Charles A. Ferguson, Uli Frauenfelder, Rochel Gelman, Diane Horgan, Peter Jordens, Aditi Lahiri, Philip Lieberman, Lise Menn, Marilyn M. Vihman, Jill G. de Villiers, and Jürgen Weissenborn for contributing comments, data, and references. I am specially indebted to Melissa Bowerman, Willem J. M. Levelt, and Dan I. Slobin for all their comments, insights, challenges, and discussion over the years.

Grateful thanks also to Margaret Amara, Center for Advanced Study in the Behavioral Sciences, Stanford, and to Karin Kastens, Max-Planck-Institut für Psycholinguistik, Nijmegen, for locating so many articles and books for me; to Laura Weeks and Susan Toubmanoff for translating articles on Russian word-formation; and to Jens Allwood, Ruth A. Berman, Werner Deutsch, Barbara Kryk, Svenka Savić, Richard Schupbach, Magdalena Smoczyńska, Ragnhild Söderbergh, Stephen Wechsler, and Alessandro Zucchi for their willingness to help me with facts and data in German, Hebrew, Italian, Polish, Russian, Serbo-Croatian, and Swedish.

Herbert H. Clark, Lila R. Gleitman, and Dan I. Slobin made detailed comments on an earlier version of this book. I have incorporated many of their suggestions and I remain indebted to them for their enthusiasm and their encouragement.

My writing was supported in part by a Fellowship from the John Simon Guggenheim Memorial Foundation, and in part by the Max-Planck-Gesellschaft. My own research has been supported by the National Science Foundation (SOC75-17126, BNS80-07349), the National Institute of Child Health and Human Development (1-R01-HD18909), the Sloan Foundation, and The Spencer Foundation. I am particularly grateful to the Center for Advanced Study in the Behavioral Sciences at Stanford for support when I originally began this project, and to Willem J. M. Levelt and Wolfgang Klein for making it possible for me to complete the writing at the Max-Planck-Institut für Psycholinguistik, Nijmegen, The Netherlands.

Finally, I have special thanks for two people: Herb, long-time reader, critic, and friend, without whom this book might never have been written, and Damon, whose words I could not have done without. This book is for them both.

Contents

<i>Acknowledgments</i>	xi
1 The lexicon: words old and new	1
The lexicon	2
Acquiring a lexicon	13
Summary	16
1 LEXICAL ACQUISITION	19
2 Early lexical development	21
First words	22
Early words and semantic fields	28
First meanings	32
Word class and word structure	38
Summary	42
3 The mapping problem	43
Ontological categories	43
Conceptual constraints	49
Lexical constraints	62
Summary	66
4 Conventionality and contrast	67
Conventionality	67
Contrast	69
Summary	83

5	Pragmatic principles and acquisition	<i>page</i> 84
	Conventionality	84
	Contrast	90
	Morphology and allomorphy	102
	Summary	108
6	Transparency and simplicity	109
	Complex words	109
	Types of mapping	111
	Transparency of meaning	115
	Simplicity of form	119
	Summary	125
7	Productivity	126
	Structure-based models of productivity	127
	Norm-based models of productivity	132
	Function-based models of productivity	133
	Contemporary preferences	135
	Productivity in acquisition	137
	Summary	140
2	CASE STUDIES OF LEXICAL INNOVATION	141
8	Words for things	143
	English	143
	Other Germanic languages	151
	Summary	159
9	More words for things	160
	Romance	160
	Slavic	165
	Two more languages	169
	Summary	175

10 Words for agents and instruments	page 177
English	178
Icelandic	182
Hebrew	186
French and Italian	190
Summary	196
11 Words for actions	198
New verbs in English	198
Other Germanic languages	205
Romance and Slavic	209
Two more languages	213
Summary	217
12 Words for undoing actions	219
Undoing in English	220
Undoing in German	229
Undoing in other languages	234
Summary	237
3 CONCLUSION	239
13 Issues for acquisition	241
Lexical acquisition	241
Production does not equal comprehension	245
Representing linguistic knowledge	251
Structure versus process	254
Coda	259
<i>Bibliography</i>	260
<i>Index of names</i>	293
<i>Index of subjects</i>	299

1 *The lexicon: words old and new*

Words make a language. They are used to talk about everything, from bee-keeping to bicycling, from navigation to international banking. They supply us with the means for everyday talk about our surroundings and activities, about people, objects, and places, about relations, properties, and states of being. We need them to communicate about events and ideas, technology, science, philosophy, and art. The stock of words speakers can draw on in a language is the lexicon. The present study is concerned with how children acquire a lexicon.

Words come first in language acquisition. Once children have some words, they begin to make generalizations about kinds of words – words for persons, places, and things, say, compared with words for actions and states. Children need words to instantiate syntactic categories, whether at the word level (noun, verb, adjective) or at the phrase level (noun phrase, verb phrase). And they must have words to realize grammatical relations like ‘subject-of’ or ‘direct-object-of,’ or mark such relations through agreement between subject and verb, say, for number and person. Without words, there would be no sound structure, no word structure, no syntax. The lexicon is central in language, and central in the acquisition of language.

The lexicon, as a result, offers a unique window on the process of acquisition for language as a whole. Although researchers have looked through this window at how children master the structure of sounds within words (the phonology of the language), and at the inflections on words (part of the morphology of a language), few have used this window to look at how children find out about, make use of, and build on words and word structure in lexical development. The present study makes such an attempt by focussing first on children’s emerging knowledge of the conventional lexicon and then on their growing knowledge of word structure reflected in their formation of new words.

What do children have to learn when they learn the words of their

language? They must learn the word forms in use among the speakers around them, and they must learn the meanings those forms carry. More precisely, for each word, they need to store all sorts of information in memory. They must store the meanings for each word, the word class it belongs to (whether it is a noun or a verb, for example), its internal structure (whether it can be broken down into smaller parts), and how it is pronounced. Before we can look at how children acquire the lexicon, therefore, we must first look at what they have to learn. The focus of the present chapter is on the nature of the lexicon.

The lexicon

The lexicon of a language is the stock of established words speakers can draw on when they speak and have recourse to in understanding what they hear. This stock is stored in memory in such a way that speakers can locate the relevant units to use in both speaking and understanding. To do this, of course, speakers have to be able to identify words either by looking them up in memory (for comprehension) or by retrieving them as appropriate forms for conveying specific meanings (for production).

Words constitute the smallest semantic units that can move around in an utterance. They can move around to form new sequences with different meanings. Compare *The man chased the dog* with *The dog chased the man*. This mobility contrasts with the immobility of morphemes inside words. Morpheme order is fixed, as in the word *chased* versus the non-word **ed-chase*,¹ or *calmly* versus **ly-calm*. The grammatical category of a form may suggest that two (or more) words have the same form. Compare the verb *open* in *Rod opened the door* or *The door opened* with the adjective *open* in *The open window* or *The door is standing open*. Occasionally, even in the same grammatical setting, a word may have such distinct meanings that one posits two (or more) distinct words, for example *bank* in *He fished from the river bank* versus *That bank is a good example of art deco*.

These expressions must be stored in memory. But just what sort of information goes into memory with each one? How is this mental lexicon organized? One can think of this vast memory store as being organized like a dictionary – a mental list of lexical items together with detailed information about each one. The lexical items can be words (*cat*, *shell*, *rowan*) or idiomatic phrases (*cook someone's goose*, *do someone in*, *go to bat*

¹ An asterisk (*) is used to mark forms as ungrammatical.

for someone, blow one's own trumpet), each with its own "entry" in the mental lexicon.

Lexical entries

Lexical entries must include at least four kinds of information about each item: (a) the meaning, (b) the syntactic form, (c) the morphological structure, and (d) the phonological shape. The lexical entry for *skier*, for instance, might be outlined as follows:

SKIER:

- (a) meaning: 'one who skis'
- (b) syntax: category Noun, count
- (c) morphology: word root + *-er*
- (d) phonology: /skiər/

The information in (a) and (b) together comprise the LEMMA and the information in (c) and (d) the FORM for a word (Levelt 1989). So the lemma and form together make up the information associated with the lexical entry for each word or phrase in the lexicon.

The meaning in a lexical entry can be characterized broadly as the conceptual information that is tagged or pointed to by the lexical item in question. If we take the perspective of a speaker planning to talk, then the meaning may be a set of conceptual conditions that must hold for a particular word or phrase to be selected. In the case of a listener, the word points to a particular piece of conceptual knowledge. The meaning specified in the lexical entry is a shorthand for the pertinent conceptual information, as in the possible glosses offered for the meaning of *skier*: 'someone who skis' or 'someone who moves over snow by means of skis.' (Notice that glosses like these do not enlighten anyone who does not already know what skis or snow are.)

The meaning in a lexical entry is linked internally to other parts of the entry. It's closely linked to the set of syntactic properties, that is, to all the aspects of structure relevant to the possible syntactic environments. In the case of *skier*, this includes the syntactic category, namely NOUN, with the further specification that it is a count noun. The meaning portion of the entry is also linked to the morphology. The noun *skier* is built from the word *ski* combined with an affix, *-er*. This affix is further specified as indicating the agent of the action denoted by the verb root *ski*, and as

attaching to a verb (or noun) to form a noun. Lastly, the meaning is linked to the phonological specification of the form in terms of the segments (/s-k-i-ə-r/), syllables (ski + er), and word stress (on the first syllable).

Lexical entries may also include information about the status of lexical items. They may indicate that an expression is dialectal, that is, not from the same dialect as the speaker's other vocabulary; that a term is characteristic of a special register or style of speech (baby talk, foreigner talk, formal speech, and so on); that it bears specific connotations, as in the differences between *statesman* and *politician*, *attempt* and *try*, or *skinny* and *slim* (Waldron 1979); or entries may contain information about usage, as for expressions like *break a leg*, *hello*, or *you're welcome*.

Lexical entries for verbs include more elaborate syntactic information than for most nouns. In addition to specifying the syntactic category as VERB, the entry also specifies the number of arguments. For a transitive verb, for example, there are two arguments, and these realize the grammatical relations of SUBJ (subject-of) and OBJ (direct-object-of). The lexical entry also indicates which roles are carried by the arguments (e.g., AGENT, PATIENT/THEME, LOCATION, and so on). So an intransitive verb like *run* would list a single argument (subject) in the syntactic portion of its entry, with the role of agent, as in *The boy runs*; transitive *read* would list two arguments, subject and direct object, with the roles agent and theme respectively, as in *The child read the book*; and transitive *put* would list three arguments (subject, direct object, and oblique) with the roles agent, theme, and location, as in *The woman put the flowers on the table*.

The morphological portion of the entry contains all the variant forms of each word. It identifies the root form (e.g., *run* or *give*), or the constituent parts of forms like *white-wash* or *compartmentalize*. In addition, for the verb *give*, it would include the forms *give*, *gives* (third singular, present tense), *gave* (past tense), and *given* (past participle). That is, this part of the lexical entry captures the intuition that the same *word* is involved in all uses of the verb *give* (*give*, *gives*, *giving*, *give*, *gave*, *given*); in all uses of the noun *horse* (*horse*, *horses*), or all uses of the pronoun *we* (*we*, *us*, *our*, *ours*). All the INFLECTED forms of a word belong to the same lexical entry.

Lexical items, then, are grouped into sets that link all the inflected forms of the same word within a single lexical entry. In a language like English, the inflections mark only the singular/plural distinction in nouns, and aspect, tense, and number in verbs. Case in English is marginal, and appears only in the subject and object forms of some pronouns (e.g., *I/me*, *he/him*, *she/her*, *they/them*). Adjectives in English can be inflected for

degree, so forms like *red*, *redder*, *reddest* would belong in the same lexical entry. In other languages, nouns (and adjectives) may be inflected for case and gender in addition to number, and verbs may mark person, gender, and number in addition to tense and aspect. The range of distinctions in each of these inflectional categories varies across languages.

The inflected forms of a word are gathered into a single lexical entry. Words that are derived from a single form, though, are generally each sufficiently specialized in meaning to have their own lexical entries. For instance, *curious* and *curiosity* have separate entries, as do *eat*, *eater*, and *eatery*; or *act*, *active*, and *action*. For derived forms, both meaning and morphology interconnect all the lexical entries that contain a particular root (e.g., *paint* in *paint*, *painter*, *painting*). There are also interconnections among lexical entries that contain the same derivational affixes (e.g., all the words with *-er*, with *-tion*, with *-ity*, or with *-ness*). These interconnections link lexical entries through meaning (for each affix), syntax (the resultant syntactic category of the derived word), and morphology. But although word meanings in the lexical entries may be related through the root in each of the derived forms, the syntax often differs considerably (e.g., for nouns versus verbs, or for adjectives versus nouns).

At the same time, some lexical items with similar meanings may show parallels in their syntax and morphology. For example, syntactically, verbs with causative meanings (e.g., *bring*, *feed*, *break*) are all transitive and so have two arguments, assigned to the grammatical relations subject and object; these arguments mark the thematic roles of agent and theme or patient. Causativity itself may also be marked in the morphology with affixes like *-ify* and *-ize* (e.g., *causativize*, *nullify*). Similarly, nouns with the meaning 'state of being ADJ' may be marked by the affix *-ness* combined with an adjective (*green*, *silky*) as in *greenness*, *silkiness*. Affixes, then, mark part of speech and added meaning for the derived word in the lexical entry.

Some meaning units are larger than words. Idioms like *have a bee in one's bonnet* (be obsessed), *go west* (get lost), *be off one's rocker* (go crazy) require their own lexical entries. With each of these, the meaning of the whole (glossed in parentheses) differs from any meaning constructible from the parts (but see Wasow, Sag, and Nunberg 1983). Compare the idiom *kick the bucket* (meaning 'die') with the non-idiomatic phrase *kick the bucket* to describe an act of kicking. The idiomatic interpretation is often restricted syntactically compared with the non-idiomatic one. One can use the idiom in the simple present, future, and past tense, but not with progressive aspect (?*The old man's kicking the bucket*), and not in the passive (**The*

bucket was kicked by the old man). Syntactic restrictions on idioms vary from one idiom to another. Some are very restricted. *Break a leg*, for example, is only used in the imperative in its idiomatic sense, but *blow one's own trumpet* (with the sense 'boast'), provided it retains *own*, can occur in a much larger range of constructions (e.g., Fraser 1974).

Word formation and innovation

The kinds of lexical entries considered so far constitute the stock of established words and phrases speakers draw on. These established terms are conventional in that all speakers in a community agree on how to use such terms. But the lexical store is not fixed in size or unchanging in membership. Words get added and lost over time. Speakers coin new words to fill gaps in the established lexicon. These coinages may be used on only one occasion (nonce uses) or may answer some need common to a larger community and eventually be added to the established lexicon. Speakers typically choose the forms for such words from existing resources. The options they draw on fall into two major classes of word-formation – COMPOUNDING and DERIVATION.

Compounds are usually divided into types according to the syntactic class of the resultant word. In English, one finds compound nouns formed from combinations of roots only, and hence often called root compounds (e.g., established *sun-rise*, *push-chair*, *dog-sled*). One also finds compound verbs (e.g., *to white-wash*, *to side-step*), and compound adjectives (e.g., *gray-eyed*). Compounds may include affixes, as in the nouns *clock-mender* and *washing-machine*. These are sometimes called synthetic compounds (in contrast to root compounds). Compound nouns, like *snow-flake*, contain a head (*-flake*) and a modifying element (*snow-*), with the head carrying number agreement, and, in other languages, case and gender as well. In compound verbs, like *to dry-clean*, the head (*-clean*) carries aspect and tense as well as any agreement for person, number, and gender. In English, the head is the rightmost member of the compound (*-flake* in *snow-flake*), and compounds generally carry primary stress on the modifier and tertiary stress on the head.

Derivations are words formed with affixation added to a word or root. Affixes (prefixes, suffixes, or infixes) may maintain or change syntactic word class (e.g., *re-* with no change in *redraw* versus *-ize* for a noun to verb change in *hospitalize*). Derivational affixes can be divided into two classes: primary or Class I affixes which typically require some modification of the root they are added to, and secondary or Class II affixes which do not