

Exemplars at work:
Theoretical arguments for
Exemplar-based Construction Grammar

A Dissertation Submitted to
The Graduate School of Letters
Keio University

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Literature

by
Masato Yoshikawa

February, 2015

Acknowledgments

There are almost countless number of people I have to thank for completing my dissertation. First of all, I am most deeply grateful to my supervisor, Professor Ippei Inoue at Keio University. He is of course a linguist, but is not an expert at the field of linguistics I am devoted to. Nevertheless, every time I made a presentation at the seminar class held by him, he always gave me insightful, suggestive and sometimes critical comments, which, clearly, contribute to the improvement of my research. Outside of the academic context, he is also a good professor, especially at a casual drinking party. He loves beer; I also love beer.

Though he is not my supervisor officially, Professor Kow Kuroda at Kyorin University can be said to be another supervisor of mine, practically speaking, so I would like to express my deepest gratitude to him. I dare say that a large part of my ideas related to the theory presented in this dissertation are based on those from his previous papers, presentations, and comments and discussions at seminars or academic meetings. He has always been inspiring me in an academic sense since I first read his paper. I have also enjoyed talking with him a hundred times (mainly over beer!), which is also academically valuable. Though he is much senior to me in terms both of age and of career, he friendly talks and even discuss things with me. In other words, he is a person of just the opposite of *authoritarianism*.

I would like to extend my thanks to Professor Yukio Tsuji at Keio University for expanding my vision, especially in the field of cognitive and neural science, as well as linguistics. I would also like to express my thanks to Professor Norimitsu Tosu, a former supervisor of mine, for inviting me to the world of linguistics when I was a sophomore (more than ten years ago!).

Other than those professors there are many who I would also like to thank. Let me thank my fiends at the English Department of Keio University, including those who have already left. Especially, in academic contexts, talks with and comments by the following friends were valuable and insightful: Professor Fuminori Nakamura, Mr. Takashi Hamada, Mr. Yoshihito Izawa, and Mr. Daisuke Nonaka. In addition, on not necessarily academic aspects, the time

spent on talking with the following people, in many cases over beer, were invaluable: Mr. Jo Thae Ho, Dr. Yukari Kato, Mr. Daichi Ishikawa, Ms. Misato Hori, Mr. Yoshihiro Sugai, Mr. Yu Nagashima, Ms. Misaki Akimoto and Ms. Saori Daiju. In addition, I would like to thank the other members of the seminar class held by Professor Inoue: Ms. Liudmila Gosteva, Ms. Sayaka Ogasawara, Ms. Fumino Horiuchi, Mr. Koichi Kosaka and Mr. Hitoshi Kawabata.

I also have several “academic friends” outside the English department of Keio, and with them I have made discussions on linguistics and other topics many times, which certainly contributed to the development of my knowledge and way of thinking. Let me thank them: Professor Noburo Saji, Professor Naoki Otani, Professor Seiji Nashio, Mr. Yukinori Kimoto, Dr. Yufuko Tkashima, Dr. Ayami Jo, Dr. Gaku Kajimaru, Professor Daisuke Yokomori, Dr. Yoshihiko Asao and Dr. Yuzo Morishita.

Further, there are two senior researchers I would like to express my thanks for inspiring me with their insightful academic works. They are: Professor Shohei Hidaka and Professor Takashi Tsukimoto.

In addition to them, my thanks go as well to a number of friends on Facebook, for they have given me hundreds of pieces of valuable information.

This may seem rather strange, but I would also like to thank the greatest drink in the world, namely, Beer. Good beers. IPAs. Weizens. Thanks, beers.

Related to this, I am also grateful to the staff at a pub located near Mita campus of Keio University, which we call *our house*, for always extending their warmest welcome to me and providing good beers. The pub is Mita 82 Ale house. Let me give a special thanks to the manager Mr. Iwakura.

Finally I would like to extend my sincere thanks to my family members. First let me thank my sister Mayumi Morisaki, my brother-in-law Ken-ichi Morisaki, my nephew Taiyo and my niece Urara for their warmest supports.

My sincerest thanks are also extended to my wife, Hiroi. I cannot write up my dissertation without her encouragement and support. She has also given spurs to me occasionally, which has been, and will probably be, the most invaluable.

I cannot end my acknowledgments without thanking a person who has long, long been supporting me, completely without expecting anything in return, namely, my mother. I couldn't thank her more.

Contents

Acknowledgments	i
1 Introduction	1
1.1 Exemplar-based Construction Grammar	1
1.2 Philosophical backgrounds	2
1.3 The organization of this dissertation	4
2 Theoretical foundations	6
2.1 Introductory notes on this chapter	6
2.2 Exemplar theory	9
2.3 The theory of grammatical constructions	25
2.4 Presenting a new theory: Exemplar-based Construction Grammar	33
2.5 Evidence for exemplar-based nature of language	38
3 Exemplar-based Construction Grammar	40
3.1 Overview	41
3.2 Background assumptions	42
3.3 Theoretical conceptions	42
3.4 Methodology	54
3.5 Advantages of EBCG over ABCG	57
3.6 Scope of the study	58
3.7 Concluding remarks	67
4 Exemplar-based accounts for English ASCs	70
4.1 Introductory notes	70
4.2 The ditransitive construction	73
4.3 The resultative construction	91

4.4	The caused-motion construction	106
4.5	The <i>way</i> construction	132
4.6	Concluding remarks	144
5	Exemplar theory meets Zipf's law	149
5.1	Zipfian distribution of rank-frequency correlation	149
5.2	Analysis of large balanced corpora	160
5.3	Analysis of child-directed speeches	192
5.4	Concluding remarks	200
6	Concluding remarks	206
6.1	Exemplar-based characterizations of constructions	206
6.2	Exemplars and Zipf's law	207
6.3	Future issues	208

List of Tables

2.1	A tiny memory space with ten exemplars	23
3.1	Components of η_l	48
4.1	The number of clauses with respect to the number of lexical arguments (Du Bois 2003:35, Table 2 [with slight modifications])	77
4.2	The significant $\langle V, N \rangle$ pairs in [VERB PRONOUN ARTICLE NOUN] on COCA	79
4.3	The frequency list of verb (forms) in the context [she . . . him something] . . .	86
4.4	X in [<i>give</i> DETERMINER X ARTICLE . . .]	90
4.5	Verbs appearing in the resultative construction s with RP <i>dead</i> and <i>to death</i> in BNC (Boas 2003:130-131)	93
4.6	The verb-RP pairs frequently used in passives	94
4.7	The verb-RP pairs frequently used in inverted order	95
4.8	The noun-RP sequences frequently found in the resultative construction	97
4.9	X in [X PRONOUN mad]	101
4.10	X in [X PRONOUN clean]	102
4.11	[X THING <i>to</i> HUMAN]	113
4.12	[X HUMAN <i>to</i> HUMAN]	113
4.13	[X THING <i>to</i> PLACE]	113
4.14	[X HUMAN <i>to</i> PLACE]	113
4.15	[X THING <i>into</i> HUMAN]	114
4.16	Words preceding [ARTICLE <i>ball inside</i>]	119
4.17	[X PRONOUN into the]	123
4.18	[X PRONOUN into \neg the]	123
4.19	X in [X POSS way PREPOSITION] on COCA	138
4.20	PMA for <i>Ann asked (him) the way</i> (Kuroda 2000:24)	144

4.21	PMAX of <i>*He walked his way to the store.</i>	144
4.22	PMAX of <i>He walked his way down the slope.</i>	145
4.23	PMAX of <i>She fought her way into the room.</i>	145
5.1	The most frequent verbs in three constructions from mothers' speech	154
5.2	Parameters and variables in various equations	158
5.3	Data summary of COCA	161
5.4	Comparison of <i>X</i> in [<i>X ARTICLE ball inside</i>] between COCA and BNC . .	163
5.5	<i>X</i> in [<i>the book X</i>]	167
5.6	<i>X</i> in [<i>book the X</i>]	168
5.7	<i>X</i> in [<i>I walked X ARTICLE</i>]	170
5.8	<i>X</i> in [<i>I X in ARTICLE</i>]	170
5.9	<i>X</i> in [<i>ARTICLE X in</i>]	173
5.10	<i>X</i> in [<i>X PRONOUN in</i>]	175
5.11	<i>X</i> in [<i>ARTICLE X into</i>]	177
5.12	<i>X</i> in [<i>X PRONOUN into</i>]	177
5.13	<i>X</i> in [<i>ARTICLE X to</i>]	178
5.14	<i>X</i> in [<i>X PRONOUN to</i>]	178
5.15	<i>X</i> in [<i>ARTICLE X to ARTICLE</i>]	180
5.16	<i>X</i> in [<i>X PRONOUN to ARTICLE</i>]	180
5.17	<i>Y</i> in [<i>Y ARTICLE X' to ARTICLE</i>]	181
5.18	<i>X</i> in [<i>X PRO₁ to ARTICLE</i>]	183
5.19	<i>X</i> in [<i>X PRO₂ to ARTICLE</i>]	183
5.20	<i>X</i> in [<i>X PRO₃ to ARTICLE</i>]	183
5.21	<i>X</i> in [<i>X them to ARTICLE</i>]	184
5.22	<i>X</i> in [<i>X me ARTICLE</i>]	186
5.23	Prepositions appearing in <i>X</i>	187
5.24	<i>X</i> in [<i>X DETERMINER NOUN to death</i>]	189
5.25	<i>X</i> in [<i>X DETERMINER head off</i>]	189
5.26	<i>X</i> in [<i>X POSS head off</i>]	191
5.27	<i>X</i> in [<i>X ARTICLE ball inside</i>]	192
5.28	<i>X</i> in [<i>X POSS way PREPOSITION</i>]	193
5.29	A data snapshot of Brown Corpus	195

5.30 Adam files	195
5.31 Adam: <i>X</i> in [PRONOUN X PRONOUN]	198
5.32 Eve: <i>X</i> in [PRONOUN X PRONOUN]	199
5.33 Sarah: <i>X</i> in [PRONOUN X PRONOUN]	199
5.34 Adam: <i>PRONOUN-PRONOUN</i> in [PRONOUN X PRONOUN]	200
5.35 Adam: <i>X</i> in [do you X]	201
5.36 Eve: <i>X</i> in [do you X]	201
5.37 Sarah: <i>X</i> in [do you X]	202
5.38 X-Y in [X Y me a/an] on COCA	204
5.39 X-Y in [X me a/an Y] on COCA	205

List of Figures

2.1	MINERVA2 (based on Figure 1 and Figure 2 in Hintzman 1986:413–414)	11
2.2	What happens when a certain feature is selectively attended (based on Figure 2 in Nosofsky 1986:42)	12
2.3	Schematic illustration of an exemplar-resonance model by Johnson	13
2.4	Episodic traces of two sentences, <i>girl who dances likes tango</i> and <i>boy likes mango</i> , with their parsings	14
2.5	Lexical network yielding word-internal morphological structure	19
2.6	“Fusion” of Caused-Motion Construction with the verb <i>sneeze</i>	27
2.7	Radical Construction Grammar representation of verbal categories	31
3.1	Alignment of syntactic elements based on semantic relations	59
3.2	A sample spectrogram encoding a voice saying “Hi.”	63
3.3	Associative relations and symbolic relations in exemplar memory	66
3.4	Phonology and morphology in exemplar memory	67
4.1	Relation network among elements	81
4.2	A content-based characterization of the ditransitive construction	83
4.3	A hierarchical network of $\langle V, N \rangle$ pairs in the ditransitive construction	83
5.1	A rank-freq distribution of <i>Brown Corpus</i>	150
5.2	A rank-freq distribution of <i>Brown Corpus</i> (log-scaled)	151
5.3	Logarithmic perception of quantity [where $k = 10$ and $S_0 = 2$]	152
5.4	Logarithmic perception of quantity with X-axis log-scaled	153
5.5	How the graph changes according to B	159
5.6	The web interface of COCA	162
5.7	X in [X ARTICLE <i>ball inside</i>]	164
5.8	Google Books Ngram Viewer	165

5.9	[ARTICLE <i>ball inside</i>] in American and British English	165
5.10	<i>X</i> in [<i>the book X</i>]	168
5.11	<i>X</i> in [<i>book the X</i>]	169
5.12	<i>X</i> in [ARTICLE <i>X in</i>]	173
5.13	<i>X</i> in [X PRONOUN <i>in</i>]	175
5.14	<i>X</i> in [X PRONOUN <i>in</i>], where <i>put</i> \notin <i>X</i>	175
5.15	<i>X</i> in [ARTICLE <i>X into</i>]	177
5.16	<i>X</i> in [X PRONOUN <i>into</i>]	177
5.17	<i>X</i> in [ARTICLE <i>X to</i>]	178
5.18	<i>X</i> in [X PRONOUN <i>to</i>]	178
5.19	<i>X</i> in [ARTICLE <i>X to</i> ARTICLE]	180
5.20	<i>X</i> in [X PRONOUN <i>to</i> ARTICLE]	180
5.21	<i>X</i> in [X PRO _{<i>i</i>} <i>to</i> ARTICLE]	183
5.22	Histogram of <i>X</i> in [X <i>me</i> ARTICLE]	186
5.23	<i>X</i> in [X <i>me</i> ARTICLE]	187
5.24	<i>X</i> in [X DETERMINER NOUN <i>to death</i>]	189
5.25	<i>X</i> in [X DETERMINER head off]	189
5.26	<i>X</i> in [X POSS <i>head off</i>]	191
5.27	<i>X</i> in [X ARTICLE <i>ball inside</i>]	192
5.28	<i>X</i> in [X POSS <i>way</i> PREPOSITION]	193
5.29	Adam: <i>X</i> in [PRONOUN X PRONOUN]	198
5.30	Eve: <i>X</i> in [PRONOUN X PRONOUN]	199
5.31	Sarah: <i>X</i> in [PRONOUN X PRONOUN]	199
5.32	Adam: <i>X</i> in [X <i>me</i> ARTICLE]	201
5.33	Eve: <i>X</i> in [X <i>me</i> ARTICLE]	201
5.34	Sarah: <i>X</i> in [X <i>me</i> ARTICLE]	202

List of Notations

[...]	a linear sequence of words (e.g, [... into pieces])
[... <i>italics</i> ...]	a lemmatized word in a sequence (e.g., [<i>make</i> it into ...])
[... CAPITALS ...]	a part of speech or some semantic/grammatical category in a sequence (e.g., [... PRONOUN into ...], [... into ARTICLE])
typewriter-faces	a program name, an URL, source codes or a corpus query

Chapter 1

Introduction

In this dissertation, a theory called *Exemplar-based Construction Grammar*, henceforth EBCG) is presented. EBCG is a theory of language which investigates the exemplar-based nature of language, focusing on the *construction effect*. The construction effect is characterized as a kind of intuition of *classification* in which we recognize something new as an instance of a thing we already know. Specifically, EBCG is aimed at explaining the mechanism of language processing by humans as a classification process with *associations* of *exemplars*, that is, individual fragments of past experience stored in mind. This chapter provide a brief overview of the theory presented in this dissertation, EBCG, and the philosophical background of it, composed of three theses.

1.1 Exemplar-based Construction Grammar

Exemplar-based Construction Grammar (EBCG) is, as the name suggests, a version of a linguistic theory called *Construction Grammar* (e.g., Croft 2001; Goldberg 1995), in which the knowledge of language is assumed to be composed of various kinds of *constructions*, i.e., stored symbols of form-meaning pairs, with varying size and degree of abstractness from morphemes such as *construct* and *-ion* to syntactic frames such as [*Subject Verb Object*]. At the same time, also as the name suggests, EBCG is a theory assuming the *exemplar-based* nature of the target phenomenon. The term *exemplar* is a hypothetical entity of our memory, long been discussed in the field of psychology, especially of cognitive psychology (Hintzman 1984; Medin & Schaffer 1978; Nosofsky 1986:e.g.,). Exemplars are considered to be the tokens of our past experience stored in mind, which are utilized when processing any newly-encountered tokens.

EBCG assumes that

- we human memorize all the exemplars of sentences previously heard/read
- when processing an input sentence we *associate* those concrete exemplars with the input;
- then we *construct* meanings (and perhaps, forms) via *integrating* or *blending* the associated exemplars and *transfer* the integrated meaning (and form) to the input.

This means that EBCG is not based on abstract entities, i.e., *constructions*, but based on concrete exemplars, hence *exemplar-based*.

Suppose, for example, we hear/see the following sentence:

(1) She kicked me a question.

This sentence would be novel and sound somewhat bizarre, but probably be interpretable. EBCG tries to pursue the *source* of interpretation, if possible, and provide empirical evidence for the process of interpretation. The source is considered to lie in a *set of exemplars* of, for example, sentences. In the case of the above example such sentences as *She asked me a question* and *She gave me a question* are the candidates for the source of interpretation. We could process the sentence based on the partially overlapping expressions previously heard, seen or even imagined.

1.2 Philosophical backgrounds

Behind the theory presented in this dissertation, there are three basic philosophical or meta-theoretical stances, namely the *constructionist thesis*, the *anti-abstractionist thesis* and the *cognitive-realist thesis*. Here the three theses are presented.

What makes the theory of grammatical constructions remarkable is that its assumption of the nature of grammar as a set of constructions. The majority of grammatical theory assumes that the grammar is a kind of *dynamic* system which is composed of *rules*, *principles* or *constraints*, from which the expressions of a certain language are generated. For example, the grammar of English is in many cases conceptualized as a set of rules yielding well-formed English sequences of linguistic elements such as phrases, clauses and sentences. The rule would include, for example, a rule by which a determiner *a* is composed with a noun *dog* and, as a result, a noun phrase *a dog* is made. Construction Grammar does not think in this way. It assumes that the noun phrase *a dog* is not composed of the determiner *a* and the noun *dog* in a bottom-up fashion, but is yielded by *instantiating* an abstract pattern or a *type* of expression,

in this case a *noun phrase*. In other words, under the theory of grammatical constructions all the expressions are *categorized* as any of the previously-set abstract constructions.

This is the *constructionist thesis*:

(2) Constructionist thesis

Expressions are not composed of or decomposed into their component parts, but *instantiate* or *are categorized as* any of abstract expression types in whole.

Even for such a large unit as a sentence we can categorize it as a certain *sentence type*, which is usually called an *argument structure* such as [Subject Verb Object₁ Object₂].

All the versions of Construction Grammar should take the constructionist thesis, and in this sense the thesis cannot be a part of unique characterization of EBCG. Assumptions of EBCG are original in terms of another stance, which is related both to the theoretical and to the methodological aspects of it, namely the *anti-abstractionist thesis*:

(3) Anti-abstractionist thesis

Both in theorizing and describing constructions they should be characterized using as concrete entities as possible.

For example the sentence presented above in (1) can be seen as an instance of what is called the *ditransitive* construction such as *He gave me a gift* and *I told him the story*, usually characterized as a highly abstract sequence of grammatical functions like [Subject Object₁ Object₂]. However, the sentence can also be analyzed based on such semi-fixed sequences as *She ... me a question* and *She kicked ... and* regarded as an instance of such lexically-specific patterns. It can, in the most extreme case, also be seen as one resembling other sentences such as *She asked me a question* and *She kicked the ball*. Which degree of specificity is suitable would depend on the target to be theorize or describe, but, as long as any meaningful characterization is available, the more specific one is preferred. In this case, perhaps, the exact sentence-based analysis sounds the best.

Here a question may arise: why should we go maximally specific? Obviously, the more general the characterization is, the more systematic and widely applicable it is. For example if the sentence (1) is characterized as a direct instance of the highly abstract structure, [Subject Verb Object₁ Object₂], then the same characterization is applicable to a wide range of expressions including *He gave me a gift*, *I told him the story* and so on. However, such a

high abstractness may lead to another question as to the *reality* of its existence: is there actually such a thing as an abstract argument structure? What makes the theory skeptical on this point is the *cognitive-realist* thesis:

(4) **Cognitive-realist thesis**

Analyses should be based on things considered as cognitively real.

This thesis leads to a number of methodological constraints such as the avoidance of using part of speech information. Part of speech would have some cognitive reality, but we cannot depend on them for analyzing our online processing of expressions because there is in many cases no concrete cue to the part of speech of a word, especially in the cases of such open-class categories as nouns and verbs.

As a corollary of the three theses, the philosophy of EBCG is obtained. Namely:

- (5) Theorize and describe expressions as instances of some types, which are characterized as concretely as possible in presumably cognitively-realistic way.

This is the reason why the current theory assumes that a sentence is processed based on associated concrete exemplars.

1.3 The organization of this dissertation

In chapter two the theoretical foundations of this dissertation are provided. First, as a model of linguistic knowledge, a group of theories called *exemplar theory* is introduced. Exemplar theory is what investigates exemplar-based nature of human memory. Second, a linguistic theory named *Construction Grammar* is introduced. Construction Grammar is a theory of language which assumes that the knowledge of language is composed of various kinds of *constructions*, i.e., stored symbols of form-meaning pairs. Third, it is argued that the two frameworks, exemplar theory and Construction should meet here in order to explain the nature of language, especially when we take the matters on *learning* and the *frequency effect* of inputs on learning.

Chapter three provides the details of the theoretical frame work, namely *Exemplar-based Construction Grammar* (EBCG). In that chapter, first, the overview of EBCG is briefly displayed (3.1), then, secondly, the background assumptions underlying EBCG are given (3.2), thirdly its theoretical conceptions are presented one by one (3.3), fourthly the methodology

of EBCG is introduced (3.4), fifthly a kind of *meta-theoretical* argument is given for showing the advantages of EBCG over the currently prevailing alternatives (3.5), and lastly the scope of the theory is manifested.

Chapter four investigates four major grammatical constructions of English, discussed in many literatures. They are 1) *the ditransitive construction*, *the resultative construction*, *the caused-motion construction* and *the way construction*, as exemplified below:

- | | |
|--|------------------------------|
| (6) a. Freddy gave me the globe. | (ditransitive construction) |
| b. Mark pushed the door open. | (resultative construction) |
| c. Remy threw the book into the water. | (caused-motion construction) |
| d. Roy made his way through the crowd. | (<i>way</i> construction) |

Specifically, all the four constructions are shown to be *explainable* based on exemplars and some surface patterns (defined in chapter 3), based on findings from previous studies and empirical data obtained from corpora.

Chapter five provides a group of quantitative case studies done to verifying EBCG's assumptions. The analyses are based on the well-known distribution of frequency called *Zipf's law* (Zipf 1935, 1949), hypothesizing words appearing at a certain position of a surface pattern shows the Zipfian distribution, that is, the frequency distribution in which the rank and the frequency are inversely proportion to each other. The hypothesis is tested with two types of data, a large-scaled balanced corpus and corpora of child-adult conversations.

Chapter six provides some general remarks on the results of researches, findings, analyses and discussions done in the preceding chapters and then notes some issues remaining to be done.

Chapter 2

Theoretical foundations

While the effects of frequency are often not noted until some degree of frequency has accumulated, there is no way for frequency to matter unless even the first occurrence of an item is noted in memory. Otherwise, how would frequency accumulate?

Bybee 2010:18

This chapter presents the framework of the theory of this dissertation with somewhat detailed review of previous studies in order to lay the theoretical foundations for the theory. Additionally, some supplementary arguments including the scope and implications of the presented theory are provided. Specifically, the organization of this chapter is as follows: 1) an introductory note on this chapter is provided, 2) previous studies advocating exemplar theory are reviewed, 3) the theory of grammar called *Construction Grammar* is introduced (e.g., Goldberg 1995, 2006; Croft 2001), 4) a new theory called *Exemplar-based Construction Grammar* is proposed, and 5) pieces of supporting evidence for the proposed theory are presented.

2.1 Introductory notes on this chapter

Before going to the main part, in this section, some introductory notes are provided in order to make clear the reasons why the assumptions are held. Specifically, this section describes in what way the two theoretical frameworks mentioned just above, that is, exemplar theories and *Construction Grammar*, are beneficial and in some cases even necessary to construct a whole new theory of grammatical construction.

2.1.1 Why exemplar-based?

Language is indeed abstract in the sense that it is a kind of *concept*; it is not any kind of physical object nor of natural phenomenon. At the same time, however, language cannot

exist without any kind of physical media such as human voice and inkblot. These media are, as it were, *products* of language. There is, in addition, yet another kind of physical media of language: a *memory trace*. Our experience of language is stored in our brain in some way and the stored experience is called a memory trace. Memory traces of language can be considered as *holders* of language.

Since so-called Chomsky's revolution, modern linguistics has been putting focus on the *cognitive* aspect of language, namely the *knowledge* of language we humans have. In this sense the latter type of physical media of language seems to have been an object of linguistic study because it is indeed a kind of our cognitive property. However, somewhat strangely, this is not the case. Modern linguistics has been interested in the *abstract* aspect, as opposed to *physical* property, of language.

In other words, our knowledge of language has been considered to be quite abstract, specifically a set of grammatical rules, categories and items such as words. This holds true whether the theory is rationalist or empiricist. The difference between the two opposites is only in the assumption about where the knowledge comes from: the former, as represented by Chomskyian version of *Generative Grammar*, assumes that the basics of our linguistic knowledge are *innate* (e.g., Chomsky 1980); with sufficient stimuli everyone can acquire his/her native language automatically. This is a *maturation* view of language acquisition, in the biological sense.

In contrast, empiricists claim that language is actually *learned* based on experience. For example, in an empiricist theory known as the *Usage-based Model* of language (e.g., Bybee 1995; Langacker 1987; Tomasello 2003), it is assumed that our linguistic knowledge is acquired through a kind of *abstraction* process called *schematization*; linguistic inputs are not mere stimuli which evokes some innate source of our knowledge, but are what basically shape our knowledge. This is a *learning* view, as opposed to a maturation view, of language acquisition.

In either view, as mentioned above, linguistic knowledge is assumed to be abstract. In this sense, the latter view, an empiricist view, can be said to be somewhat, say, *half-baked*; it can be more radical if it goes to extremes, that is, assuming our linguistic knowledge as *concrete*. This assumption may sound unrealistic, but in fact it is the fundamental idea of what is called *exemplar theory*.¹⁾

Exemplar theory is a model of human cognition which considers that our knowledge and

¹⁾Note here that to be precise, an exemplar-based model does not suppose that our knowledge is completely concrete. Related topics will be presented in the following sections.

cognitive processing are based on individual *exemplars*, that is, individual fragments of our experience (e.g., Hintzman 1984; Cf. Medin & Schaffer 1978). If this view is applied to a linguistic theory, the theory will assume that our linguistic knowledge is based on a set of exemplars. It can be said that this is a truly empiricist view of language.

The important point here is that whether one takes rationalist or empiricist view, it can only be a hypothesis; in other words, for now, there is no firm evidence to confirm which view is empirically right. Under such a circumstance, it is beneficial to hold as few assumptions as possible because if the view taken were to be proved to be incorrect, all the assumption should be discarded. Consequently, as to the basic assumption on the knowledge of language, we should reduce theoretical assumption as much as possible.

In this connection, empiricist view can be said as advantageous in that it can incorporate empirical data as supporting evidence without any limitations. The rationalist view, in contrast, cannot do the same thing because of its innateness hypothesis: it only thinks of input data as triggering factors which evokes the innate knowledge, so it should additionally stipulate some other mechanisms which shape our linguistic knowledge and convert input stimuli to specific forms of the knowledge. In Generative Grammar theory, the two mechanisms are actually provided: the abstract mental grammar which is considered as a set of rules, principles or constraints known as *Universal Grammar* and what is called Language Acquisition Device (LAD), respectively.

There is, as is evident, no need for such additional components of linguistic knowledge if we take an empiricist view. From such a philosophy of science point of view, therefore, other things being equal, an empiricist approach is concluded to be better than a rationalist one. Then the next question is: how empiricist should we be?

The answer is: to the extent possible. This is a kind of an *empiricist enterprise*, embodied as an exemplar-based theory of language. In order to construct an innovative theory to solve some unaddressed but important problems, this kind of extremeness is considered to be necessary, because the problem to be solved itself may in most cases be somewhat novel and hence hard to solve by already-known measures. Of course empirical data should be consulted, but as long as it is *falsifiable* and *not falsified* by data, the theory should be free from data; more specifically, its assumptions and logics can be organized independently of empirical data.

2.1.2 Why constructionist?

Generally, as seen in the previous chapter, there are two possible ways of modeling sentence processing: *parsing* and *type judgment*. The former assumes that in order to process a sentence we should decompose it into some smaller units such as words and then reconstruct what appears to be an original structure, namely the *syntactic structure* of it. In other words sentence processing is assumed to be a process of, somewhat metaphorically, *dissection*, in the sense that it is based on the idea that to understand something is to know its inner structure.

On the other hand, type-judgment model assumes that in order to process a sentence it is enough for us to know what the sentence is in its entirety. Its inner structure is considered to be what we do not necessarily know. Sentence processing under this model can be seen as a process of *categorization*: we can categorize some citrus as an orange only with its external, superficial features and do not need to know its internal structure or component parts.

Simply, almost all of our daily cognitive activities can be seen as versions of the latter type of processing. Analyzing internal structure is, say, a kind of *metacognitive* activity, which often needs some expert knowledge or retrospective contemplation. Language processing is certainly a daily activity in that we humans do it everyday almost unconsciously, without any expert knowledge.²⁾ In fact the majority of English speakers (perhaps including some expert linguists) cannot provide proper analyses of such pervasive sentences as the following, but they never fail to categorize them as such and such types of sentences they already know:

- (7) a. It rains cats and dogs.
- b. The bigger they come, the harder they fall.

It is, therefore, enough for us to recognize and understand what the confronting sentence means without knowing its internal structure. In other words it is enough to know what *type* of a sentence it is. Consequently, in view of the reality of our daily cognition, the type judgment model can be said as, at least, preferable in the sense that it can reflect our daily cognitive activity more directly than the parsing model.

2.2 Exemplar theory

In the days of Chomsky's revolution, the idea of exemplar-based processing did not exist. The basic idea underlying any implementations of exemplar theory are highly simple and,

²⁾Of course we can say that the knowledge of a language is a kind of expert knowledge and, for example, call an English speaker as an expert of the English language, but this line of argument hardly makes sense.

probably because of the very simplicity, it had not been treated in a systematic and principled way as a model of human cognitive processing. Even now, exemplar theory cannot be said popular in cognitive science. This, however, does not mean the theory is of no effect. The fact may be just the opposite.

The basic idea behind exemplar theory can be described as follows: every single experience is stored in mind and then recycled to process newly confronted experiences. When you first meet a person unknown before, *A*, you may feel that the person (*A*) resembles some others you already know, such as *B*, *C*, and *D*. In this case *A* is a newly-processed token based on your past experiences. After that, you meet *A* for the second time and this time you recognize *A* utilizing your past experience of *A*. Now *A* is no longer a target to be processed but, in contrast, a part of sources from which you process a new piece of information.

This may sound too simple to explain a full variety of human cognitive activities. Below the details of several exemplar models are presented to show how the idea behind exemplar-theory works, and, in addition some remarks on the key concepts of exemplar theory are provided.

2.2.1 Brief history and overview

The idea of exemplar theory can be traced to the paper by Medin & Schaffer (1978), which presents a model of our conceptual representation in mind and the process of human categorical judgment based on the conceptual model. After Medin & Schaffer (1978), a number of exemplar-theoretic models were proposed in 1980s, such as Hintzman (1984) and ?. Most of them are in the field of cognitive psychology and aim at explaining categorization behaviors through simulation. More specifically they explain how humans judge a novel item as a member of some category based on the similarity between the item and stored exemplars.

In Hintzman's (1984) model an exemplar is coded as a point of multi-dimensional space, or a feature vector, whose values are either -1 (negative), 0 (missing), or $+1$ (positive). His model named MINERVA2 computes similarity as a degree of activation using the inner product between a target item called a *probe* and an exemplar and then the activation value is multiplied by the exemplar vector. After that each of the feature values is summed to one and averaged, resulting in a new vector, called an *echo*. The echo gives the probe a content lacking in it including features coding category labels, which is considered as the process of category judgment (see Fig. 2.1).

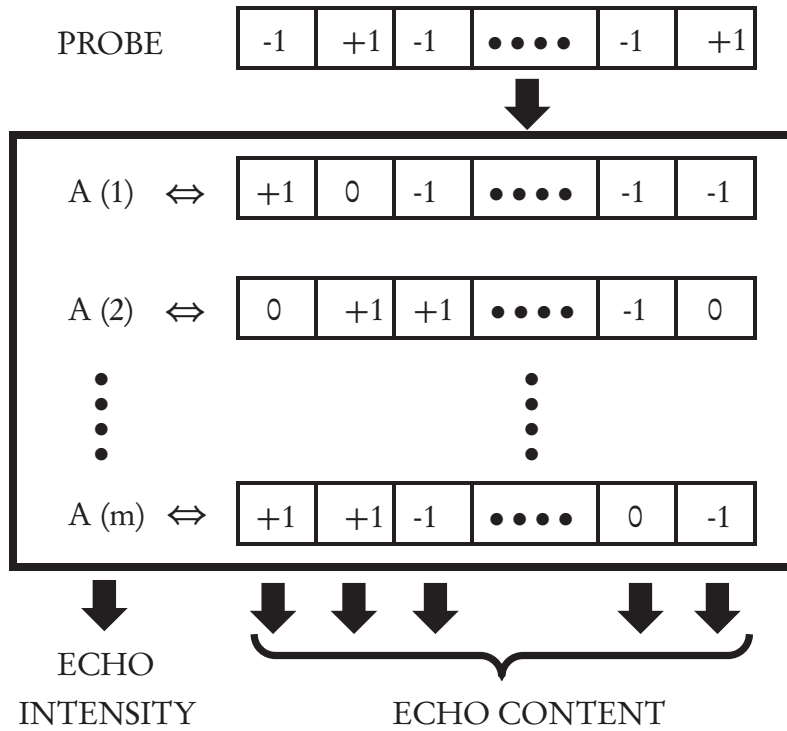


Figure 2.1: MINERVA2 (based on Figure 1 and Figure 2 in Hintzman 1986:413–414)

A model presented by Nosofsky (1986) named *generalized context model* (GCM) employs almost the same computational method to explain human categorization behavior. However, Nosofsky (1986) models not only the process of categorization but also of *identification* in a unified way. He shows the two process, categorization and identification, can be modeled by the same mechanism if an idea of *selective attention* is properly incorporated. If we selectively attend a certain sets of features (e.g., shape, size and color) when recognizing something, we tend to ignore some specificities differentiating individual tokens and, inversely, to augment some differences between a group of tokens and another, hence the thing being recognized is judged as a member of a category sharing the same values of selectively attended features.

This identification-categorization relation is diagrammatically shown by Nosofsky (1986:42, Figure2), as depicted in Fig. 2.2.1. Without selective attention the three features, shape, size and color, are treated equally, represented as the same distance among the eight corners in the upper diagram of Fig. 2.2.1. However, if a certain feature, in this case a color, is selectively attended, the distance changes: the axes of shape and size shrinks and that of color stretches, as shown in the lower diagram of Fig. 2.2.1.

There are also exemplar models for category learning, not for category judgment, such as Kruschke (1992), which presents a model named *ALCOVE* (standing for *attention learning*

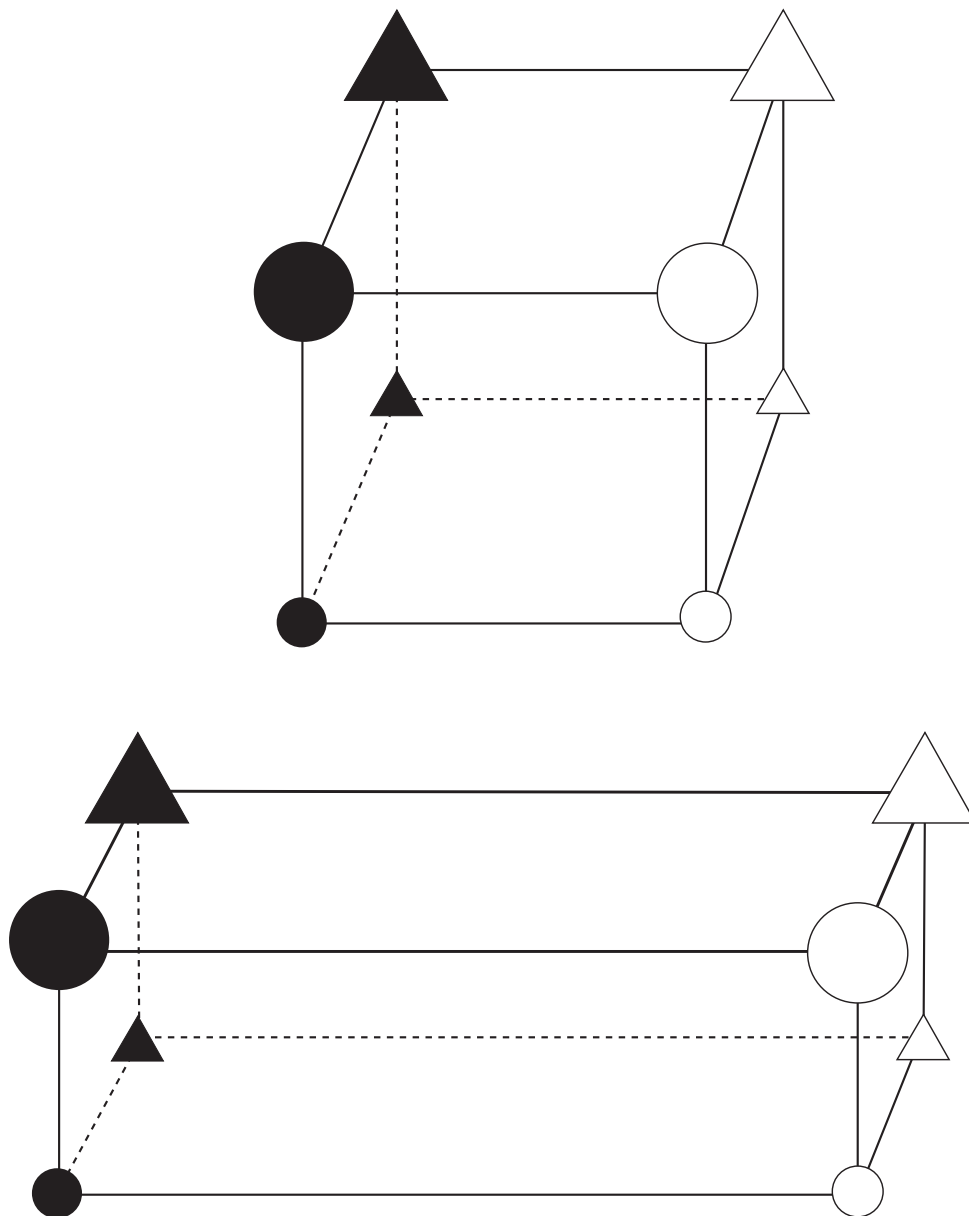


Figure 2.2: What happens when a certain feature is selectively attended (based on Figure 2 in Nosofsky 1986:42)

covering map). Kruschke (1992) applies GCM to the model for category learning, combining it with some learning models. ALCOVE is a kind of *connectionist model*, a model designed to imitate the behavior of human neural network (e.g., Rumelhart et al. 1986; McClelland et al. 1986). For convenience sake the detailed description of the architecture of ALCOVE is omitted, but a brief sketch is provided: it can be characterized as a dynamic version of GCM by Nosofsky (1986), in the sense that while GCM, as well as MINERVA2, starts with the state in which the values of multidimensional vector are already set, ALCOVE models the process of setting the values, represented as the *connection weight* of links connecting exemplar nodes

with feature nodes and category nodes in a connectionist network.

As for language, there are not so many models based on exemplars. Of them phonetic and phonological ones are famous, such as Johnson (1997) and Pierrehumbert (2001), because auditory processing can be discussed within the framework of categorical judgment. These models assume that every specific and concrete auditory experience is stored in memory with some kind of labels such as phoneme-like alphabetical ones and the sequence of them, e.g., a word or a morpheme, and the auditory details themselves are utilized when processing input data (see Fig. 2.3).

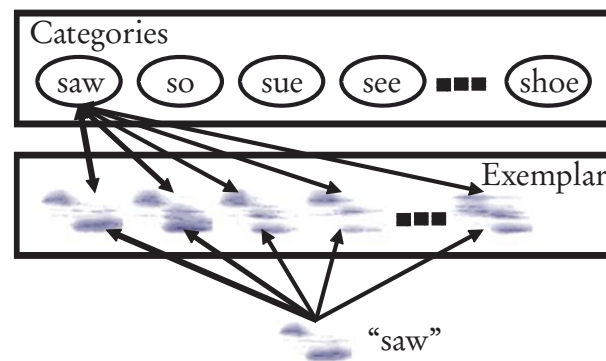


Figure 2.3: Schematic illustration of an exemplar-resonance model by Johnson (adapted from Fig. 5 in Johnson 2006:493 with slight modifications)

When it comes to syntax or semantics, there is almost no established model based on exemplars. Notable exceptions are works by Rens Bod and his colleagues such as Bod (2009) and Borensztajn & Zuidema (2011). They present a model which can be characterized as an exemplar-based model of syntax, coupled with a framework of sentence processing called *Data-Oriented Parsing* (DOP). Borensztajn & Zuidema’s (2011) model is called *Episodic Grammar*, which assumes that every sentence is processed based on the *history* of previous experiences or episodes of derivation. Derivation in their model is the process in which the internal structure of a sentence is analyzed, resulting in providing a tree structure with syntactic labels such as N (= Noun) and VP (= Verb Phrase).

Figure 2.2.1 briefly shows what the memory for sentences assumed in Episodic Grammar are like. The rectangles and triangles represent *treelets*, fragments of syntactic parse, which correspond to rewrite rules in syntax (e.g., $S \rightarrow NP VP$). The syntactic categories on the top of the shapes are, therefore, the left members of the rule and the categories at the bottom are the right members. For example, the rectangle at the bottom left corner represents a rewrite rule $[N \rightarrow \textit{tango}]$ and the triangle at the top of the figure represents $[S \rightarrow NP VP]$.

The comma-combined integers, n, m , put at the middle of each shapes are the indices of the treelets, where n , the former number, represents the index of a sentence and m indicates the order of the parse. The dashed arrows connecting treelets the actual traces of syntactic parsing memory.

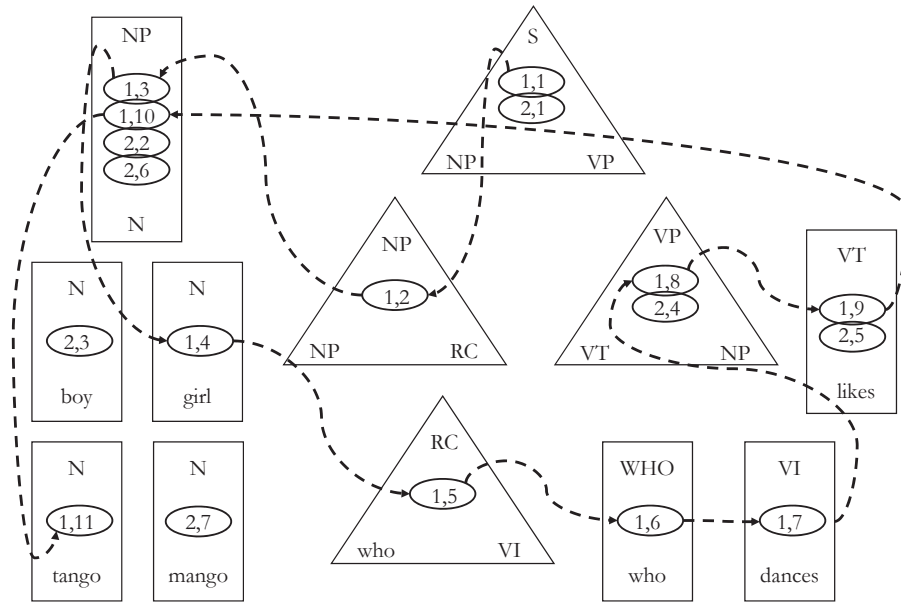


Figure 2.4: Episodic traces of two sentences, *girl who dances likes tango* and *boy likes mango*, with their parsings

(Borensztajn & Zuidema 2011:508, Figure 1)

This model is a probabilistic model which utilizes probabilities of derivations given a sequence of words, i.e., a sentence, and the most probable derivation is selected as “the best parse” and applied to the sentence. In this respect the model is somewhat different from phonetic/phonological ones given above.

The common feature shared by those models is the co-existence and interaction of exemplars and abstract properties or labels. The models are indeed exemplar-based, but exemplars alone do not suffice. Input stimuli are first processed based on concrete properties which navigate to abstract properties or labels and hence categorizations or analyses result.

2.2.2 More on exemplar models of language

In order to see how exemplar theory works as a model of language processing, it can be beneficial to go into greater detail on the exemplar models of language. Below the four types of exemplar-theoretic studies on language are provided. They are: experimental studies on

phonological memory representation, a simulation model of speech production, a model for language acquisition and representation called *Usage-based Model*, and an integrated model of language learning, language and representation called *Hierarchical Prediction Network*.

Phonological memory representation

In the context of exemplar theory, it is assumed that our the memory for phonological information is represented in a highly distributed manner, with a large variety of features including high-dimensional sound properties, voice properties and phonological category labels. This is called the “rich memory” assumption (e.g., Port 2007a). The assumption is supported by a number of experimental studies.

Palmeri et al. (1993) claimed that in memorizing a spoken word highly concrete features such as speaker’s voice property were also stored, according to the result of experiments. In the experiments, subjects were exposed to a series of words from loudspeaker and asked to judge whether a word provided last was equivalent to a word provided first. The same subject listened to several sets of words with varying number of intervening words between the first and the last. The different subjects were exposed to the word sets with different number of speaker who pronounced them; that is, the number of speakers differed between subjects.

As a result, the performance was significantly high if the first and the final words were produced by one and the same person than produced different ones, and, although the performance got worse as the number of talkers increased, the difference in performance between the “same speaker” condition and the “different speaker” condition was stable; that is, the performance in the “same speaker” condition always exceeded that in the “different speaker” condition by almost the same amount. This result strongly suggests that the process of memory recall is greatly constrained by concrete sound property such as speaker’s voice property and therefore the structure of stored memory has highly concrete character.

Goldinger (1996) also conducted very similar experiments. Though the detail of them is omitted, the result of his experiments showed that speaker’s properties were retained for a day and perceptual identification for a week. This result indicates the possibility that concrete properties are stored even in long-term memory, which suggests that our memory structure is based on highly concrete information.

A problem of phonological representation is a problem of *categorization*. That is to say, the psychological problem in phonology is how we can categorize several different pieces of sounds into one category as a phoneme. Exemplar-based theory of phonology is inevitably

faced with this problem, because if our linguistic memory is based on exemplars, we must generalize them into abstract representation in some way. Otherwise, we cannot identify the phonologically same sound and hence fail to recognize abstract sound categories such as words, which results in communication breakdown.

Exemplar theorists therefore should address the problem. As a solution to this problem, non-exemplar theorists advocate a process known as *normalization* in terms of acoustic properties (reviewed in Johnson 1997:145-146) or via abstract lexical representation (e.g., Jackson & Morton 1984). This normalization-based approach, however, was found to be implausible through a large number of experimental studies (Goldinger 1996:1166-1168). There are quite a lot of evidence which shows that we humans utilize specific voice properties in processing speech.

Speech production

There is also another kind of variability, namely variability within the same person: one and the same person speaks differently from situation to situation. Of course, within-speaker difference can be bundled together in terms of, for example, speaker's voice property. In fact, Goldinger (1996) revealed that voices with highly similar sound property were differentiated if produced by different person. Therefore, within-speaker variability can be said to be subtle. In other words, the voice of one and the same speaker is somewhat stable. Then, how is this stability attained? To clarify this, we need to examine the process of speech production.

Pierrehumbert (2001) performed simulation experiments in which the process of word production in speech was simulated. She modeled exemplar-based theory of phonological memory to simulate speech production. In the model, it was assumed that perceptual information was stored as exemplar and each exemplar had parameters which were stored in "parameter space," different from "exemplar space" in which exemplars themselves were stored (Pierrehumbert 2001:140-144). Parameters were considered to have granularity, and hence some two exemplars were recognized as the same though not exactly the same.

She conducted three simulation experiments. In the first simulation, the performance was gradually getting worse as a number of iteration increased, which was not compatible with human performance: our pronunciation becomes stable as we grow (Lee et al. 1999, referred to by Pierrehumbert 2001:149). The second experiment, with a modification to the model of the first simulation (added a *Systematic Bias* to the first model), also ended with the undesirable performance (as in the first simulation). In the third simulation, the model was

further modified and elaborated. So far, the model selected the target exemplar in such a way that among the exemplars associated with the activated label, that is, a sound to produce, one exemplar was randomly selected. “Instead, a target location in the exemplar cloud is selected at random, and the exemplars in the neighborhood of this location all contribute to the production plan” (Pierrehumbert 2001:150) in the third simulation. At this time, the model successfully simulated speaker’s gradual “entrenchment,” that is, decrease of variance.

Usage-based Model

As introduced in 2.1.1, there is an empiricist theory of language called *Usage-based Model* (e.g., Bybee 1995; Langacker 1987; Tomasello 2003). Usage-based Model (UBM) assumes experience-based learning as the process of language acquisition and in this respect it can be seen as a kind of exemplar-based learning model such as Kruschke (1992). However, as discussed in 2.1.1, the currently prevailing version of the model can hardly be regarded as an example of exemplar-based theory of language learning. The crucial difference between exemplar theory and UBM is the treatment of individual exemplars. In UBM, concrete exemplars are only the sources of information obtained through the process of learning. Each exemplar is assumed to be accumulated and contributes to the learning of some abstract unite, pattern or structure the exemplar instantiates.³⁾

However, there are at least two notable exceptions: Bybee (2010) and Taylor (2012). Bybee (2010:14), for example, explicitly argues the importance of the rich exemplar memory, as well as abstract and generalized pieces of linguistic knowledge. Specifically, she mentions the process of construction learning and its mental representation in an exemplar-based manner, taking the construction called the *resultative construction* as an example, referring to the corpus-based study by Boas (2003). The resultative construction includes sentences with the verb *drive*, as seen in the following example (Bybee 2010:26, all the examples are retrieved from *British National Corpus*):

- (8) a. It drives me crazy.
- b. they drive you mad
- c. that drives me mad
- d. A slow-witted girl drove him mad.

³⁾For example, Tomasello (2003) almost invariantly uses the word *exemplar* meaning almost the same as *instance*, as seen in the expression like “an exemplar of some more abstract construction or constructions” (Tomasello 2003:106)

e. It drove the producer mad.

The construction, as Boas (2003) points out, has a highly limited productivity in that it only allows a small number of *resultative predicates* such as *crazy* and *mad* and the predicates have quite strong preferences to cooccurring verbs such as *drive*. In other words, the construction can be characterized with several combinations of a verb or verbs and a predicate or predicates (see 4.3 on more details description of this construction and discussion about it). Moreover, the variable slot in which an object appears is also characterized as having a strong preference of occurrence: it is “most commonly a pronoun” (Bybee 2010:27). Similarly, even the subject position, which is far less restricted and far more variable than the object position, may be somewhat preferred to the two pronouns, *that* and *it*, as seen in (8) (Bybee 2010:27).

In this way, a grammatical construction can be described as a set of concrete exemplars with prototypical lexical sequences such as *It drives me me crazy*. In other words, constructions are considered as based on a set of exemplars, not vice versa. The point here is that for learners, the exemplar is, at least at first, not an instance of an abstract grammatical construction (in this case the resultative construction) because no one knows it is, but only a *singleton* exemplar. The construction should be seen to be *constructed* gradually in the course of language learning base on exemplars, as all the studies advocating Usage-based Model assume (e.g., Tomasello 2003, as will be seen in 2.3.3), so we, as analyzers, cannot utilize the information of each exemplar being an instance of any abstract construction to be learned.

Joan Bybee has long argued for this item-based nature of linguistic representation in mind, described as an *associative network* (e.g., Bybee 1995). Concrete items are connected with their partial overlaps of form and meaning and consist of a vast network, from which an abstract schema or structure such as a morpheme and a grammatical construction emerges. Figure 8 illustrates a local piece of lexical network yielding morphological structure. It should be mentioned that, however, the network itself does not represent a network of exemplars, but of labels or features, which are abstract and general. In this sense all of her works on the associative network cannot be regarded as the studies embodying exemplar theory. Nevertheless, the idea of associative network has much similarity to that of exemplar theory and has a possibility of being extended to an exemplar model.

Taylor (2012) also makes arguments for exemplar theory of language. As the title of the book, *The mental corpus*, shows, he regards our linguistic knowledge as a kind of large set of concrete examples of text and speech, resembling a corpus. This is exactly an idea of exemplar

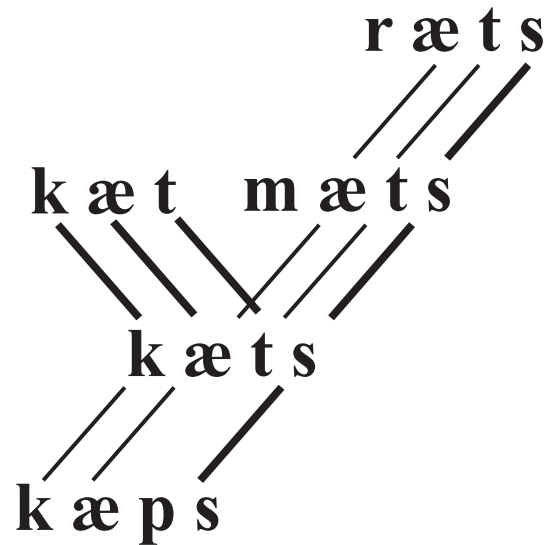


Figure 2.5: Lexical network yielding word-internal morphological structure
(from Bybee 1995:429, FIG. 1)

theory. He proclaims that “memory traces, linked by patterns of similarity and related by emergent generalization, are all there is to knowing a language.” Rather radically, he, for example, denies the invariability of word meaning, claiming that the meaning of a word is only the matter of its use. Instead he emphasizes the importance of linguistic *context* as a determiner of the sense of a word (Taylor 2012:Chapter 9, especially in pp. 241-244). Context is unique to each occurrence of a word and in that sense the context-based characterization can be seen as quite exemplar-theoretic. In fact the pioneering work of exemplar theory by Medin & Schaffer (1978) presents a exemplar-theoretic model which is named *context theory*.

Hierarchical Prediction Network

We cannot avoid mentioning the study by Borensztajn (2011) when we talking about exemplar theory of language. Borensztajn (2011) presents an integrated model of language, called *Hierarchical Prediction Network* (HPN), which covers language acquisition, on-line processing of a sentence and the mental representation of linguistic concepts including grammatical categories and semantic features. He applies a general learning model called *Memory-Prediction Framework* (Hawkins & Blakeslee 2004) to language acquisition. His processing model is, as mentioned above, an exemplar-based parsing named *Episodic Grammar* (Borensztajn & Zuidema 2011). As for the model of internal representation of linguistic knowledge, he adopts an model combining semantic and episodic memory (Cf. Tulving 1972) in order for the model

optimal to his learning and processing models.

The Memory-Prediction Framework (MPF) is a model of the behavior of cortex, which enables humans to store a huge amount of information and, based on the information, to *predict* what happens next, which are, as the authors say, what we call “thinking” and, when sensory input is given from the outside, “perception” (Hawkins & Blakeslee 2004:104). In other words, MPF regards our cognitive activities as a variety of prediction based on stored information in cortex.

Among several assumptions and arguments MPF holds, Borensztajn (2011:12) emphasizes its *hierarchical* aspect. The motivation for this is that he intends to use the framework as a model of language learning, which needs some kind of hierarchical organization of items as seen in category formation and sentence structures. The hierarchical aspect of MPF can be described, in short, as a model of abstraction with decreasing resolution of information.

However, he points out three shortcomings of MPF when it is applied to language processing (Borensztajn 2011:20-21). Of the three the most important deficiency for the argument here is that MPF does not have the place for episodic memories, which both Borensztajn (2011) and the theory of this dissertation think play an important role in language processing. In fact Borensztajn (2011) supplements his model with an additional processing architecture called Episodic Grammar, as introduced above. Moreover, it is in this respect that Borensztajn’s (2011) theory deserves mention in the context of exemplar theory of language.

He models the process of *memory consolidation* in which individual episodic experiences are gradually (re-)arranged and (re-)organized into a systematic and partially generalized network (Borensztajn 2011:31-34). This, however, does not mean that episodic memories are gradually eliminated during the process of consolidation. Instead, they become interconnected more and more densely, with their shared semantic features (Borensztajn 2011:32-33).

2.2.3 Key concepts

As described above, there are several models and studies based on the exemplar-theoretic idea, not all of which are compatible with each other. Approaches and assumptions may differ among them, but we can find some shared concepts in them and those concepts are considered to be important building blocks of exemplar theory. Here, therefore, the four concepts which are at least partly shared by the exemplar models are described in some detail. The concepts are: *exemplar*, *association*, *learning*, and *abstraction* and generalization .

Exemplar

It is beneficial to make clear what an exemplar is in exemplar theory. As seen previously, in the computational models such as Hintzman (1986) and Nosofsky (1986) an exemplar is represented as a multidimensional feature vector, which means it is regarded as a set of features. This characterization is, however, rather technical one, and perhaps there is an implicit assumption behind it on what an exemplar is like: namely, an exemplar is, as it were, only a *placeholder* of features. In other words it is just a trace of an experience and the content of it is a set of parameters.

It can be said that this view is shared by the standard idea on what an episodic memory is like and how episodic memories are organized. As Borensztajn (2011:32) summarizes, episodic memories are assumed to be “constructed as pointers that bind together items stored in semantic memory.” “Items in semantic memory” here correspond to the features or parameters assumed in the computational models of exemplar theory, so the characterization of an episode is also a contentless trace without any parametric features combined with it.

This way of characterizing an exemplar (or an episode) leads to the distinction between specific items in memory and an exemplar. The two concepts are frequently confused but in fact they are totally different. A memory of a specific person *A* is a specific item, and a memory of an event in which you met *A* yesterday is an exemplar. No exemplar occurs twice; they are unique. Specific items have their contents on their own, not mere traces of experiences.

In the theory presented in this dissertation, however, an exemplar is defined somewhat differently. The details will be provided in the next chapter, specifically in 3.3.1 and subsequent subsections. In those sections, moreover, a distinction between an exemplar and an episode is also introduced (especially in 3.3.4). In short, an exemplar in the theory is not treated as a mere trace or placeholder, but an experience of some symbolic sequence, and an episode is regarded as a sequential grouping of several exemplars experienced in a limited but not so small amount of time.

Association

Exemplars are only the sources of information stored in mind. We need, therefore, a procedure to manipulate them. The procedure assumed in exemplar models is, as occasionally mentioned, *association*. Association is thought to be an unconscious and automatic process

provoked by any input stimulus, connecting the stimulus and stored memories based on similarity between them. For example Hintzman (1986) implements *associative recall* in his model as a basic components of recognition memory.

Learning

Learning in exemplar theory can be seen as a process of memory consolidation, as Borensztajn (2011) characterizes. Memory consolidation is “the gradual construction of a relational semantic network out of episodic memories” (Borensztajn 2011:164). This process can be seen as somewhat similar to the category learning process implemented in ALCOVE (Kruschke 1992), in that the connection weights between exemplars and categories are considered to be tuned gradually. The point here is that the nature of exemplars, or episodes, itself is assumed not to change during the course of learning. Exemplars are unique and unchangeable because they are physical traces. Learning in exemplar theory should, therefore, be a process of *organizing* stored exemplars in order to optimize them for effective processing.

Let us take a simple example. Suppose a child first encounters a furry animal moves around a room. The encounter is memorized by the child as a unique event or episode, with several pieces of information such as the place where the event happens, the sound the animal makes, the visual image of the animal, the voice from mother being heard during the event, and so on. At this moment the child may not know what the animal is, but is probably able to recognize the animal as an individual entity which autonomously moves, that is, an animate being, thanks to some innate knowledge.

Later the child undergoes similar experiences over and over again, and every single event is stored as an exemplar, with various kinds of information, a large part of which are shared among the experiences. During this process the child gradually learns, for example, that the animal is a single individual, what the animal is called, how the animal behaves, how the child should behave to the animal, how mother behaves to the animal, and so on, based on the properties attached to each encounter with the animal. Some properties such as perceptual features may play the role in earlier stages, but more conceptual features such as “happened in the room” can be used if the child has already learned the concept of “the room,” a specific room the child spends most of time, probably a living room of the child’s house. Features and interrelations between the features are also learned.

This is the process of learning in exemplar theory. The unique events of encountering the animal never change, but can be reanalyzed retrospectively, if the children later finds

commonalities between/among any parts of them.

Abstraction and generalization

Due to the very nature of exemplar-based-ness, exemplar theory is quite often misunderstood as assuming models without generalization and/or abstraction. This is, however, not the case. It can be said that exemplar theory tries to implement generalization without any explicit generalized components. In other words, generalization can be done only with exemplars accompanied with a large number of features. Generalization is the process of grouping specific items together and making inferences about some unknown aspects of the group. Categorical judgment is the typical case of generalization, because a category is a group of items and the category of any new input is unknown information to be inferred. No one knows what a category is exactly like; it can only be inferred based on past experiences of a members of the category. In this sense, exemplar theory surely involves the process of generalization.

The abstraction process, on the other hand, can not be said to be addressed by all the exemplar models. A representative model which deals directly with the process is MINERVA2 by Hintzman (1986), introduced in the previous section. As the title of the paper includes the phrase “schema abstraction,” the model is not only for categorization but also for an abstraction. It produces a vector, *echo*, which has values computed based on the similarity between an input stimulus, a *probe*, and each exemplar. The echo has a character which can be seen as a result of abstraction from all the exemplar. This may be confusing, so below a sample case of abstraction with a miniaturized version of MINERVA2 is provided.

Table 2.1: A tiny memory space with ten exemplars

	1	2	3	4
$E(1)$	-1	+1	-1	+1
$E(2)$	-1	+1	-1	+1
$E(3)$	+1	+1	0	+1
$E(4)$	-1	0	0	-1
$E(5)$	+1	0	+1	-1
$E(6)$	-1	0	+1	+1
$E(7)$	0	-1	-1	+1
$E(8)$	-1	-1	-1	0
$E(9)$	0	-1	-1	+1
$E(10)$	+1	-1	-1	+1

Suppose a memory space with four different features and ten exemplars. The values of each feature in those exemplars are set as described in Table 2.1, each of whose columns cor-

responds to a feature, indexed with an integer j where $1 < j < 4$, and rows to an exemplar, represented as $E(i)$ where $1 < i < 10$. Here a new input, P , with values $[0, +1, +1, -1]$ is given to the model. First the similarity between each exemplar, $E(i)$, and the probe (represented as $S(i)$) is computed by an equation (Hintzman 1986:413):

$$S(i) = (1/N_R) \sum_{j=1}^n P(j)T(i, j) \quad (2.1)$$

where N_R is the number of features relevant to the current processing, meaning that “the number for which either $P(j)$ or $T(i, j)$ is nonzero” (Hintzman 1986:413), $P(j)$ represents the j th feature value of the probe and $T(i, j)$ represents the j th feature value of i th exemplar. For example, $S(1)$ is calculated as follows:

$$S(1) = (1/4)(0 \times -1 + 1 \times 1 + 1 \times -1 + 1 \times -1) = \frac{(0 + 1 - 1 - 1)}{4} = -0.25$$

$S(i)$ for each exemplar is presented in Table 6.3.

$E(i)$	$S(i)$	$E(i)$	$S(i)$
$E(1)$	-0.25	$E(6)$	0
$E(2)$	-0.25	$E(7)$	-1
$E(3)$	0	$E(8)$	-0.5
$E(4)$	0.25	$E(9)$	-1
$E(5)$	0.5	$E(10)$	-0.75

The similarity score $S(i)$ is, then, cubed and the cubed similarity becomes the degree of activation of each exemplar, $A(i)$ (Hintzman 1986:413):

$$A(i) = S(i)^3 \quad (2.2)$$

Finally, each feature value of the echo, $C(j)$, is computed as follows (Hintzman 1986:414):

$$C(j) = \sum_{i=1}^m A(i)T(i, j) \quad (2.3)$$

For example, $C(1)$ is calculated as follows:

$$C(1) = (-0.25^3 \times -1) + (-0.25^3 \times -1) + \dots + (0.5^3 \times 1) = -0.15625$$

The echo vector has values $[-0.15625, 2.515625, 2.703125, -2.59375]$. The values of this vector, however, does not fall within the range of -1 to $+1$, so the values are normalized by multiplying each value by $g = 1/\max[C(j)]$ (Hintzman 1986:416), in this case $1/2.703125$. The normalized vector is: $[-0.058, 0.931, 1, -0.96]$ (rounded to three decimal place).

The echo vector, $[-0.058, 0.931, 1, -0.96]$, highly resembles the probe, $[0, +1, +1, -1]$, but not the same as it. In addition, the vector is generated by the similarity-based interaction between the probe and each exemplar. This means that the input stimulus causes the stored exemplar to generate a pseudo-exemplar similar to the probe. The pseudo-exemplar, moreover, has values which are neither -1 , 0 nor $+1$, which can be interpreted as some degree of abstractness. In this way, exemplar models can implement the process of abstraction without assuming abstracted entities such as schemas, but with on-line computation based on similarity between input stimulus and stored exemplars.

2.3 The theory of grammatical constructions

Construction Grammar (CG) is a theory of grammar which can be characterized as a monostatal, declarative approach to grammar. By monostatal it is meant that the theory only assumes one level of structural representation, that is, it does not assume any kind of derivation. By declarative it is meant that the components of the theory are static, that is, no rule-like operations are assumed.

This character marks a sharp contrast with that of rather classical theories of grammar, especially a family of theories known as *Generative Grammar*, which assume that the elementary part of grammar is a set of rules which operate some kinds of items such as syntactic categories (e.g., *Sentence*, *Verb Phrase*, and *Noun*) and words.⁴⁾ Most of those traditional theories regard a verb as the core or the *head* of a sentence, that is, a sentence is, say, considered to be an instance of a verb. Therefore, the structure of a sentence is assumed to be determined by the properties of the verb contained by the sentence.

⁴⁾There was an intermediate case between rule-based and template-based theory of grammar, known as *Government and Binding Theory* (Chomsky 1993). The theory assumes that the structure of a sentence is composed of syntactic schemas called *\bar{X} schemas* (Jackendoff 1977), which provide a template of canonical phrases such as a verb phrase and a noun phrase. The theory, however, also has derivation rules which transform the templates into somewhat different configurations, hence intermediate between rule-based and template-based.

For example, the structures of following sentences are thought to be determined by the verbs, *sleep*, *break*, *hit*, and *make*:

- (9) a. Louis slept well. (Intransitive)
 b. Judy hit the ball. (Transitive)
 c. Freddy gave me the globe. (Ditransitive)
 d. Sarah made him happy. (Causative)

This assumption is made probably because in those theories the meaning and structure of a sentence is thought to come only from the component parts, that is, words included in it and syntactic categories, and not from elsewhere.

However, there are many verbs which can be used in various structures, or more precisely, *argument structures*. For example, the verb *kick* can be used in at least following eight argument structures (Goldberg 1995:11):

- (10) a. Pat kicked the wall.
 b. Pat kicked Bob black and blue.
 c. Pat kicked the football into the stadium.
 d. Pat kicked at the football.
 e. Pat kicked his foot against the chair.
 f. Pat kicked Bob the football.
 g. The horse kicks.
 h. Pat kicked his way out of the operating room.

In order to explain this fact, verb-centered theories should assume that verbs of this kind have as many meanings and potential structures as they occur in different structures. This, obviously, leads to a circularity (Goldberg 1995:10–12).

What is more, the meanings of the whole sentence vary systematically according to the argument structure in which the verb is embedded. This fact suggests that, contrary to the assumption of the generative theories, there can be sources of the meaning of a sentence other than component parts of the sentence.

In this connection, CG assumes that an argument structure also has its own meaning and is a kind of *construction*, a pair of form and meaning. Construction *C* is defined as follows:

- (11) *C* is a CONSTRUCTION iff_{def} *C* is a form-meaning pair, $\langle F_i, S_i \rangle$ such that some aspect of F_i or some aspect of S_i is not strictly predictable from *C*'s component parts or from

other previously established constructions. (Goldberg 1995:4, with slight modifications)

This allows us to assume that the meaning of a verb differs according to argument structures it occurs in because, although the verb is or can be monosemous, structures themselves have their own meanings. In other words, argument structures are assumed to be *objects* in their own rights, not composed or constructed from smaller parts such as words in it.

Further, CG enables us to account for somewhat anomalous or novel uses of verbs such as follows:

(12) He sneezed the napkin off the table. (Goldberg 1995:9)

In this example the verb *sneeze* is unusually used transitively taking an object “the napkin.” In order to account for this fact verb-centered theory should assume the verb *sneeze* has a subsense and a substructure of its own which enable it to be used transitively, but this is hardly the case. CG, in contrast, can easily explain this: the semantic and syntactic properties of the verb *sneeze* can be *fused* with that of the argument structure construction, namely *Caused-Motion Construction* whose structure is [Subject Verb Object Oblique] and meaning, [X causes Y to move Z] (Goldberg 1995:5).

Fig. 2.6 graphically displays how the argument structure construction and the verb are fused. Argument structure constructions and verbs are thought to have their own sets of *semantic roles* such as {cause, goal, theme} and {sneezer}. Fusion succeeds when and only when the roles specified by the construction and the verb correspond each other and are fully compatible (Goldberg 1995:50–52).

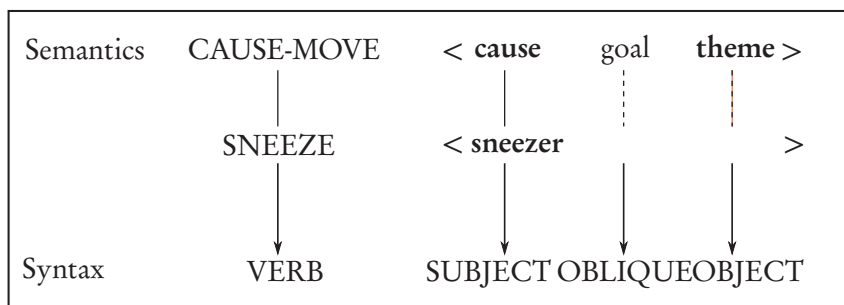


Figure 2.6: “Fusion” of Caused-Motion Construction with the verb *sneeze*
(adapted from Figure 2.10 in Goldberg 1995:54 with modifications)

2.3.1 Versions of Construction Grammar

The above depicted theorization of grammatical constructions is a representative one by Goldberg (1995), but there are some other versions of Construction Grammar. In this section, then, several versions of Construction Grammar including Goldberg's (1995) are described in some details. They are: the construction theory of the earliest stage, sometimes called *Berkeley Construction Grammar* (Sag et al. 2012:2), Goldberg's (1995) version of Construction Grammar, and the theory called *Radical Construction Grammar* by William Croft (Croft 2001).

The beginning of Construction Grammar

The basic idea of Construction Grammar was proposed by Lakoff, Fillmore, Kay, and collaborators. Lakoff (1987) analyzed English existential there-construction (e.g., *There's a fly in my soup*) based on the idea of grammatical constructions, which are defined as "direct pairings of parameters of form with parameters of meaning" (Lakoff 1987:464).

Lakoff considers that grammatical constructions formed radial categories, in which members are linked by family resemblance. A radial category is composed of relatively small number of central members, i.e., *prototypical* members, and larger number of non-central members with varying centrality of the category.

Fillmore et al. (1988) investigated English idiomatic constructions, specifically, the conjunction *let alone*. They claimed that the existence of irregular idioms, i.e., non-compositional and sometimes syntactically deviant complex units with conventional meanings, would refute any standard syntactic theories in the tradition of Generative Grammar, and therefore a new idea of grammar was needed in whose "proper units [...] are more similar to the notion of construction in traditional and pedagogical grammars than to that of rule in most versions of generative grammar" (Fillmore et al. 1988:501). By construction they meant a conventional unit which specifies "not only syntactic, but also lexical, semantic, and pragmatic information" (ibid.).

Incidentally, Fillmore and Kay's theory later somewhat changed its character by incorporating the formal syntactic theory called Head-driven Phrase Structure Grammar (e.g., Pollard & Sag 1994). Their theory is called *Sign-based Construction Grammar* (SBCG: Boas & Sag 2012). The term *sign* used in the name of theory denotes a form-meaning unit; that is, sign in Saussurian sense. SBCG is characterized as a unification-based theory of construction grammar. Unification is a process of integration in which two items with some overlapping

elements, specifically, several (typed-)features, are “unified” into one based on the overlaps.⁵⁾ In terms of this unification-based nature and some other, SBCG can be said to be quite similar to the theory proposed in this dissertation, that is, *Exemplar-based Construction Grammar*.

Goldberg’s version of Construction Grammar

The idea that grammar is composed of numerous conventional units which unfold both formal and semantic information, i.e., *constructions*, was adopted by Adele Goldberg, and Construction Grammar theory achieved a theoretical progress by her (in Goldberg 1995). Goldberg (1995) analyzed abstract kind of grammatical constructions: *argument structure constructions*. She remarked that “argument structure constructions are a special subclass of constructions that provides the basic means of clausal expression in a language” (Goldberg 1995:3)

The argument structure constructions Goldberg discussed included English ditransitive constructions, caused-motion constructions, resultative constructions, intransitive-motion constructions, and conative constructions (Goldberg 1995:3–4). Examples and structural descriptions are show below:

- | | | | |
|---------|---|-------------------------|-----------------------------|
| (13) a. | Ditransitive | X CAUSES Y TO RECEIVE Z | Subj V Obj Obj ₂ |
| | (e.g., Pat faxed Bill the letter.) | | |
| b. | Caused Motion | X CAUSES Y TO MOVE Z | Subj V Obj Obl |
| | (e.g., Pat sneezed the napkin off the table.) | | |
| c. | Resultative | X CAUSES Y TO BECOME Z | Subj V Obj Xcomp |
| | (e.g., She kissed him unconscious.) | | |
| d. | Intrans. Motion | X MOVES Y | Subj V Obl |
| | (e.g., The fly buzzed into the room.) | | |
| e. | Conative | X DIRECTS ACTION AT Y | Subj V Obl |
| | (e.g., Sam kicked at Bill.) | | |

(Goldberg 1995: 3-4, with a slight modification)

She stressed the importance of these constructions because there are “systematic differences in meaning between the same verb in different constructions” (Goldberg 1995:4). For example, when used in simple transitive sentences, such verbs as *throw*, *cast*, and *toss* typically denote a kind of event in which someone releases something from his/her hands to somewhere, while when used in ditransitive sentences (e.g., *She threw me the ball*) they represent

⁵⁾As for unification-based theory of Construction Grammar, see also Kay & Fillmore (1999)

some kind of transfer of something from one person to another. She attributed the source of differences to the meaning the constructions have. That is to say, each construction, e.g., Ditransitive construction, has its own meaning, e.g., “transfer” of something from one person to another.

As seen in (11), Goldberg (1995:4) defines construction as “a form-meaning pair $\langle F_i, S_i \rangle$ such that some aspect of F_i or some aspect of S_i is not strictly predictable from C ’s component parts or from other previously established constructions.” According to the definition, morphemes are construction because their component part, i.e., phonemes, cannot predict the meaning of them. In contrast, basic phrases such as red dog are not constructions, because the meanings of them are predictable from the component part, i.e., words such as red and dog.

Construction Grammar assumes that constructions form the global network with varying abstractness; constructions may be lexical, phrasal, clausal, super-clausal, and so on. Lakoff (1987) proposed the radial category structure of constructions in analyzing there-constructions, as mentioned above. Goldberg further developed the notion of construction network by elaborating the links connecting constructions. She proposed four types of inheritance links: polysemy links, metaphorical extension links, subparts links and instance links (Goldberg 1995:74-81).

Polysemy links connect several construction subtypes within the same polysemous construction. For example, Goldberg remarked that English ditransitive construction has various senses and they are connected by polysemy links (Goldberg 1995:75).⁶⁾ Metaphorical links connect two (or more) different constructions. For example, resultative construction and caused-motion construction were claimed to be linked by metaphorical link and the former derived from the latter (Goldberg 1995: Chapter 8). The details of the latter two links, subparts links and instance links, are omitted here.⁷⁾

Radical Construction Grammar

As reviewed so far, Goldberg ’ s version of Construction Grammar has highly elaborated character in its theoretical assumptions and explanatory devices. However, her theory leaves a little to be desired. There is yet another version of Construction Grammar: William Croft’s *Radical Construction Grammar* (RCG: Croft 2001, 2005). RCG is considered to be preferable

⁶⁾This assumption, however, has some problems. Related discussion will be provided in 3.5.2.

⁷⁾The former link, i.e., the subparts like, will be mentioned in 4.4.

to Goldbergian Construction Grammar in some respects. In this section we review Croft's (2005) contrastive argument for the advantage of his radical version of construction grammar, which gives us a preferable idea of constructions in language.

citetcroft05 labeled Construction Grammar of Fillmore & Kay's and Goldberg's as "vanilla construction grammar" (Croft 2005:273), as contrasted with his "radical" Construction Grammar. In short, the contrasted points between the two versions of Construction Grammar were as follows (VCG denotes vanilla construction grammar and RCG denotes radical construction grammar):

- (14) a. VCG: Assuming grammatical constructions as one of the components of grammar
RCG: Assuming grammatical constructions as primitives of grammar
- b. VCG: Containing construction-independent elements such as Subj, Obj, Verb (as seen in (13)) as formal representation of constructions.
RCG: Only containing construction and its component parts as formal representation of constructions.
- c. VCG: Being neutral as to language-universality of constructions
RCG: Claiming all the constructions are language specific

(From the argument in (Croft 2005:267-277))

It can be said that the thesis of RCG shown in the first contrast, (14a), renders RCG radical. VCG, as Croft (2005) remarks, only describe "part-whole relationship of complex constructions to the units that make them up" (Croft 2005:276). Therefore, the existence of non complex atomic elements such as *noun*, *verb*, *subject*, and *object* is presupposed. However, this presupposition is problematic: Croft claims that all the parts of speech such as noun or verb and syntactic relations such as subject and object are construction-specific. In other words, we cannot define these elements without referring to constructions.

In Radical Construction Grammar, however, the existence of general categories such as verb is not denied (Croft 2005:284). For example, Croft explains that the general category verb is justified by "the occurrence of verb category in another construction, namely the morphological construction of tense-agreement (TA) inflection" (ibid), which is represented as "taxonomically super ordinate category to the Intransitive Verb category, Transitive Verb category, and other verbal categories" (ibid, see Figure 14).

This claim can be supported in view of the result of experiments which examine the mechanism of category-based abstraction in the studies of *Artificial Language Learning* (ALL) (e.g.,

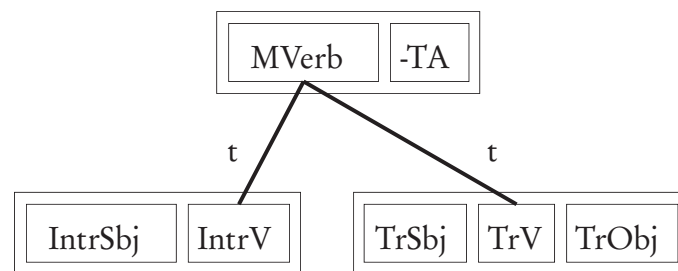


Figure 2.7: Radical Construction Grammar representation of verbal categories
 (From Figure1 in Croft 2005: 285 [*IntrSbj*: Subject in Intransitive construction;
IntrV: Verb in Intransitive construction; *TrSbj*: Subject in Transitive construction;
TrV: Verb in Transitive construction; *TrObj*: Object in Transitive construction; *t*: Taxonomic link])

Mintz 2002; Gerken et al. 2005). In these experiments, infants or adults were exposed to specific frame-based patterns (e.g., *bool* __ *jiv* in “bool nex jiv”, “bool kwob jiv”, “bool zich jiv” and “bool pren jiv”) as stimuli and succeeded in learning abstract category. This claim is also confirmed by the study examining child directed speeches in corpora from *CHILDES*, a database collecting transcriptions of infant-adult conversations (MacWhinney 2000), were analyzed (Mintz 2003). In addition, (human) language acquisition study such as Lieven et al. (1997) also reports that infants’ early speech contained several specific frame-based patterns in noun phrases (determiner + noun).

From this claim the second thesis shown in (14b) follows. If we cannot employ elements such as *subject* and *object*, we have to formalize constructions in construction-specific terms. As displayed in Figure 14, Croft does not use the general label but use construction-specific term such as *IntrSbj*, *TrObj*, and so on. This way of characterization of grammatical structure is highly compatible with usage-based theory of linguistic structure (Tomasello 2003:162-163; see also the section below (2.3.3)).

The final thesis shown in (14c) is derived from the evidence from typological point of view. Typologically, we can not find any universal kind of construction (like a universal passive; Croft 2005:277, 303-309; see also Givón 2002:22-27).

2.3.2 On the definition of *construction*

As mentioned in (11), construction *C* was defined as a form-meaning pair whose meaning and form cannot be strictly predicted by its component parts. However, the definition of construction was later modified or extended to include kinds of *predictable* units of form-meaning pairs, if they are frequently enough and assumable to be *stored* in memory (Goldberg 2006:5).⁸⁾ Here the key of construction-ness is switched from the unpredictability or *noncompositionality* to *storage* or *memory*, which can be seen as the matter of our cognition and learning.

However, this characterization does not make sense in the framework of exemplar theory. In the exemplar theory framework, our memory is considered as quite rich, in that almost all the experiences are assumed to be stored in memory (Bybee 2010:24). This topic will be discussed in detail in the next chapter, especially in 3.3.6 and 3.3.7. Further, a related topic will be discussed in 2.4.3.

2.3.3 On acquisition and learning of constructions

As briefly described in the previous section, there are several studies on the learning process of grammatical constructions. The most influential work on this topic is Tomasello (2003), which presents a usage-based theory of language acquisition. He assumes a gradual process of acquisition beginning with rote-learning of fixed and concrete expressions with some functional property and then developing into variable and abstract patterns. At around the age of three, a number of *abstract constructions* are considered to become learnable based on previously learned lexically-specific syntactic patterns, mainly the constructions called *verb-island constructions*, which are partially generalized verb-centered frames such as *Throw* __ and *Put* __ *in* __ (Tomasello 2003:Chapter 5).

The abstract constructions discussed by Tomasello (2003) include the argument structure constructions such as the ditransitive construction and the resultative construction. More importantly, the cognitive representation of the constructions are assumed in the same way as in Radical Construction Grammar (Tomasello 2003:162-163 and *passim*). This strongly suggests that a bottom-up approach to language learning is highly compatible with the radical version of construction theory. Item-based nature of the learning of grammatical construc-

⁸⁾Goldberg (2006:5) remarks: “[...] patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency [...].”

tions naturally lead to the construction-specific properties even with those of abstract types, i.e., argument structure constructions, which may probably mean that Radical Construction Grammar is also best compatible with the exemplar-based approach to the grammatical constructions, presented below.

As for construction learning, Goldberg has also conducted a number of studies on language acquisition, mainly focusing on the nature of generalization by children Goldberg (2006:e.g.,). Most of her works on acquisition are experiment-based, in which she shows the paring of syntactic patterns such as [Subject Object Verb] and eventual meaning such as [X appears on Y] can be learned by children. The studies on construction learning can also be found in the field of second language acquisition. Ellis & Ferreira-Junior (2009), for example, presents statistical analyses of construction learning by adult English-learners' speech (see 5.1.2 for more details).

2.4 Presenting a new theory: Exemplar-based Construction Grammar

So far the two theoretical frameworks, exemplar theory and Construction Grammar, has been reviewed in some detail. When it comes to the model of human language processing and representation, both of the two frameworks cannot be seen as sufficient. The insufficiency of the former framework has been already pointed out, that is: there are only few exemplar-theoretic studies on syntactic aspects of language which are systematic and general. As for the latter, it can be pointed out that the currently prevailing version of construction grammar only deals with the aspect of *representation* and that of processing is not argued about, at least systematically (this deficiency will be discussed in more detail in 2.4.4).

Fortunately, the insufficiency on one theory can be supplemented with the other. Exemplar theory is a theory of memory representation and cognitive processing on that in general, not of a language in particular. Construction Grammar is, on the other hand, a linguistic theory whose primal concern is not such psychological aspects as human memory and cognitive process. What is needed here is, therefore, to integrate the two frameworks and construct a new theory.

Here a new theory is presented, named *Exemplar-based Construction Grammar*, EBCG for short. EBCG is, as the name shows, an exemplar-theoretic version of Construction Grammar. The basic architecture of the theory will be given in the next chapter, but in the followings a

brief introduction of the theory is presented, with some meta-theoretical arguments for it.

2.4.1 Preliminaries

EBCG is proposed in order to overcome deficiencies found in the prevailing versions of Construction Grammar, when regarded as psychologically realistic models of language. This section therefore presents several problems in Construction Grammar from psycholinguistic perspectives. The problems are related to the assumption for mental representation of constructions and its learning and processing.

2.4.2 Verb/Construction dichotomy

Goldberg's version of Construction Grammar assumes that a sentence is composed by fusing an existing argument structure construction such as the ditransitive construction and a verb. This means that argument structure constructions and verbs are considered to be two distinctive entities, though both have a status of being constructions (e.g., Goldberg 1995:50-52). However, this way of conception, which can be called the *verb/construction dichotomy* (Cf. Croft 2003), would probably be a fallacious one.

Simply, it is dubious that we process newly-confronted sentences based separately on abstract, lexeme-free constructions such as the ditransitive construction and verbs such as *give*. If we see or hear, for example, the sentence *He gave me a gift*, we can most probably process it in whole, without decomposing it into two parts, [Subject Verb Object₁ Object₂] and *give* and then fusing the two into one. Of course there do exist cases where the decomposition-fusion process seems more plausible, especially in the case of a sentence with novel verbal usage in a certain construction, as seen in the *sneeze* example of the caused-motion construction (12). However, even in those cases, there is no guarantee that we utilize the two composing parts, an argument structure construction and a predicate verb, to process a sentence. It is also possible to think that the lexically-specific sequences such as *off the table* do some job to specify the structure and the meaning of the whole sentence (see 4.4.1 for related arguments), and the verb *sneeze* is only one member of those lexical or super-lexical sources, without any specialized privileged properties.

Also on the descriptive aspects, the problem of verb/construction dichotomy is pointed out. For example, ? argues that the English ditransitive construction, as one example, should be described in a verb-class-specific way, because the construction has a number of subsenses

each of which is used with verbs in a verb class specific to the subsense (e.g., verbs of giving such as *give*, verbs of refusal such as *refuse* and *deny*, and verbs of permission such as *permit* and *allow*). Constructions represented in a verb-class-specific way are called *verb-class-specific constructions* (Croft 2003:58). The arguments for verb-class-specific constructions are applied to other constructions such as the caused-motion construction (Iwata 2008) under the name of *lexical-constructional approach*.

However, even if we assume the verb-class-specific constructions, problems are not completely solved. There, at least, still remain the following questions:

- (15) a. Where do the verb-class come from?
- b. How many verb-classes are there?
- c. How are the verb-classes organized?

The problems here would be, ultimately, those of, say, *meaning determinacy*. As will be discussed in the next two chapters (3.5.2 and 4.2.2), the problem related to the ditransitive construction pointed out by Croft (2003) is of the polysemous character of the construction: if the construction is polysemous and its subsenses can be attributed to specific verb class, it would be better to divide it into several verb-class-specific constructions because of descriptive adequacy. In other words, verb-class-specific constructions are advantageous compared to abstract argument structure constructions in terms of the specificity of meaning.

From this the following principle is obtained: the more specific the description is, the better it is. Of course, specificity stands on the sacrifice of generality, but, in most cases it is unknown that how much generality is needed to adequately describe a phenomenon. Therefore, the only possible approach to meet the adequacy is going as specific as possible. This means that the ultimate version of construction description should be *exemplar-based*. As will be shown in Chapter 4, especially in 4.2.2, the polysemous character of the ditransitive construction can be described in a bottom-up fashion, based on the clusters of exemplars.

2.4.3 Frequency effects and learning

In the studies on the learning of grammatical constructions such as Tomasello (2003) and Goldberg (2006), mainly under the framework of the *Usage-based Model*, it is emphasized that the input frequency is of primary importance to learn new constructions or make new generalizations, as seen in 2.2.2. However, it can be pointed out that the frequency-based approach to construction learning has one crucial problem, which is called the *beginning para-*

dox (Yoshikawa 2009). Namely, the frequency-based account for construction learning cannot explain how the learning begins.

In the frequency-based theory it is usually assumed that for us to learn a certain construction or any other linguistic material it is necessary to experience the target item enough number of times, which is the *threshold* of learning; we cannot learn anything without repeated encounter of it. However, logically speaking, this assumption does not work. Suppose, for example, we need ten times to learn a novel linguistic item such as a word, and at some stage of our language development, we encounter a word *dog* for the first time. At this moment we cannot learn the word because we have not encountered it more than ten times, so we will probably forget the encountering. Then we encounter the same word again, but since we do not remember that we have already encountered the word, the second encounter is not recognized as such. In this way, the frequency count of a word is ever be accumulated, hence the impossibility of learning it (Cf. Arnon & Snider 2010, Bybee 2010:18, as quoted in the epigraph of this chapter).

One possible solution is to remember certain examples which would be important in preparation for the future generalization. However, we cannot know in advance which example is important to remember for future generalization, so the only approach is to remember all the example we encounter. This is the very assumption of exemplar theory.

2.4.4 Relationship between models for representation and processing

Here it should be pointed out that in the Construction Grammar literature there is almost no mention of matters on *representation* of construction. More precisely, almost no one argues about how constructions are retained in our memory. At the same time, more importantly, there are only a few studies saying something about *on-line processing* of construction; most studies do not concern about how a stream of items is recognized as an instance of some construction (notable exceptions are: Jurafsky 1992, Ellis & Ferreira-Junior 2009 and arguments of *Sign-based Construction Grammar* such as Boas & Sag 2012, as will be discussed in ??)

This is problematic in the sense that even if constructions can be described as pairings of abstract syntactic frames (e.g., [Subject Verb Object₁ Object₂]) and meanings encoding certain scenes or abstract relations (e.g., X CAUSES Y TO RECEIVE Z) and the descriptions are identified with pieces of our knowledge, the descriptions themselves cannot explain how

we can access the knowledge in the course of sentence processing. In other words, those descriptions do not provide any cues to the *linking* from a concrete sequence to be processed when encountered such as *He gave me a gift* to the target abstract construction such as the paring ([Subject Verb Object₁ Object₂], [X CAUSES Y TO RECEIVE Z]). Some may say that the clue would be the verb *give* or the sequence *gave me a ... and* this would probably be the case, but how could this be explained in a *principled*, non-ad-hoc way?

As the next chapter presents, the exemplar-based theory of grammatical constructions aims to attain exactly this goal, that is, to explain the process of online linking between a concrete sequence and an abstract knowledge based on concrete cues available in the sequence to be processed. If we assume that our knowledge of constructions are based on previously encountered concrete exemplars such as {*She gave me a book, I gave him the present, He asked me a question, ...*}, it is possible to appeal to a kind of *similarity judgment* between stored exemplars and a newly-encounter sentence being processed.

In other words, if construction theory is conceptualized under an exemplar-based framework, we can deal with both the representation and the processing aspects of constructions, at a time. How constructions are assumed to be represented in mind under the framework has not been specified yet, but will soon be described in detail in the next chapter, especially in 3.3.6 and 3.3.7.

2.5 Evidence for exemplar-based nature of language

The marriage of exemplar theory and construction grammar can be an innovative framework of language processing and representation. However, for now, it almost lacks empirical evidence showing the exemplar-based view of construction is correct. Here, therefore, some pieces of indirect evidence are provided in order to complement the lack of empirical validity of the current theory. The evidence is three-fold: on 1) the formulaicity of language, 2) the *Idiom Principle* and 3) the memory for word sequences.

If language is processed and represented in an exemplar-theoretic manner, the natural way to handle linguistic elements in memory is more like a *holistic* one than *analytic*. Words are rarely provided isolatedly; they usually appear in a sequence. Most of the linguistic inputs we gain are, therefore, multi-word sequences. It is reasonable to assume that, unless our perception of linguistic input is optimally designed to process inputs word by word, the unit of

language attached to individual exemplars is not a word, but a multi-word sequence. This assumption is also supported by the fact that the intended meaning of a word in context can hardly be identified without its linguistic context, that is, words surrounding it. In other words, most words need a process called *word sense disambiguation* in order to properly understand the meaning of a single occurrence of any word.

If this is the case, the two corollaries are obtained. One is about the output aspect of language and the other is about the input aspect. As to the output aspect, if our exemplar memories are usually connected with multi-word sequences, language production should usually be done not in a compositional way, but in, say, a *formulaic* way. Being compositional means that a sequence is produced by composing smaller units such as words into a larger one such as a sentence. This is a view taken by the majority of linguistic theory. but this view cannot be seen as compatible with the assumed nature of linguistic exemplars as discussed above. Simply put, compositional way to produce linguistic expressions is *costly*, in that in order to do so we first need to segment smaller units from holistic sequences and then recombine them into a new sequence. The alternative way which seems less costly is to produce new sequences only by partially *editing* existing expressions. This leads to the *conservativeness* of linguist production. In fact, as Alison Wray argues (e.g., Wray 2002, 2007; Wray & Grace 2007), formulaicity is one of remarkable features of language. By one estimation it is said that as much as seventy percent of our utterances are formulaic (Wray 2002:1).

As for the input aspect, it is predicted that we tend to interpret language in as holistic a way as possible. If exemplar theory is correct, linguistic inputs should be best and most effectively interpreted when the input sequence is exactly the same as any of already stored sequences attached to exemplars. The effectiveness, or the cost, of interpretation is assumed to decrease as the gap between the input and stored exemplars is widened. This tendency is actually confirmed by Sinclair (1991) based on observations of vast amount of corpus data, presented as a principle of linguistic interpretation, named the *Idiom Principle*, saying, roughly, other things being equal the idiomatic interpretation is preferred.

In addition, there is an experimental study directly examining the multi-sequential nature of linguist information attached to exemplar memories. Arnon & Snider (2010) find that four-word sequences such as *don't have to worry* and *don't have to wait* elicit difference in response time according to the frequency of the whole sequence when asked to judge whether they are possible English expressions or not, which suggests that we store the exact sequences

composing of four words and the number of encountering them directly effects how fast we can access the stored sequences. The difference in response time, incidentally, can be attributed neither to the frequencies of component words such as *worry* and *wait* nor to those of any bi- or tri-grams such as *to worry*, *to wait*, *have to worry* and *have to wait*.

Chapter 3

Exemplar-based Construction Grammar

The human cortex is particularly large and therefore has a massive memory capacity. It is constantly predicting what you will see, hear, and feel, mostly in ways you are unconscious of. These predictions are our thoughts, and, when combined with sensory input, they are our perceptions. I call this view of the brain the memory-prediction framework of intelligence.

Hawkins & Blakeslee (2004: 104)

In the previous chapter it was argued that this is the time for the theory of construction to meet exemplar theory. The reason is two-fold: 1) it is suitable for the theory to explain frequency effects and matters on learning; 2) as the other side of the first reason, exemplar theory is considered highly effective for human sentence processing. This argument is made in terms of the *cognitive-realist* thesis.

This chapter, in turn, provides some details of the theory presented in this dissertation, that is, *Exemplar-based Construction Grammar* (henceforth EBCG), under the name of *anti-abstractionist thesis* and the *constructionist thesis*. That is to say, the arguments provided in this chapter are largely theoretical, in that they are based on hypotheses constructions and hypothetical reasoning, rather than empirical data.

Specifically, first, the overview of EBCG is briefly displayed (3.1), then, secondly, the background assumptions underlying EBCG are given (3.2), thirdly its theoretical conceptions are presented one by one (3.3), fourthly the methodology of EBCG is introduced (3.4), and lastly a kind of *meta-theoretical* argument is given for showing the advantages of EBCG over the currently prevailing alternatives (3.5).

3.1 Overview

EBCG is, as the name shows, a version of a linguistic theory called *Construction Grammar* (e.g., Goldberg 1995), especially of a radical one advocated by William Croft (e.g., Croft 2001, 2005). As seen in the previous chapter, in Construction Grammar the knowledge of language is assumed to be composed of various kinds of *constructions*, i.e., stored symbols of form-meaning pairs, with varying size and degree of abstractness from morphemes to syntactic frames, which are considered to be connected into a vast network. The grammar of language is, therefore, considered to be built on the objects, i.e., constructions, as opposed to sets of derivation rules or constraints.

This assumption underlying the theory is named the *constructionist* thesis and therefore Construction Grammar can be seen as a constructionist theory. It can be said that, however, the assumption may only be one possible way of embodying constructionism. In fact, as will be seen in this chapter, EBCG embodies constructionism in a quite different way from that of the prevailing theories of construction.

Our intuition for construction can be described as something like a “this is that” feeling, which means that for a newly encountered expression we *recognize* it as an instance of something previously experienced, i.e., a construction we already know. It is true that this kind of intuition can be explained by assuming the existence of a construction in mind which gives us the intuition. However, there is another possibility to explain the intuition, namely, the assumption that the intuition is only an *effect* caused by other sources than the construction. EBCG pursues the very alternative possibility, the detail of which will be given in 3.3.6 and 3.3.7. Specifically, EBCG assumes that the intuition comes from the *association* of a set of *exemplars* stored in mind. Below a brief illustration is provided.

EBCG assumes that 1) we human memorize all the exemplars of sentences previously heard/read; 2) when processing an input sentence we *associate* those concrete exemplars with the input; 3) then we *construct* meanings (and perhaps, forms) via *integrating* or *blending* the associated exemplars and *transfer* the integrated meaning (and form) to the input. This means that EBCG is not based on abstract entities, i.e., *constructions*, but based on concrete exemplars, hence *exemplar-based*.

Taking a simple example of comprehension, suppose a novel sentence *s* such as (16) is given.

(16) She skyped me a photo.

The sentence s is considered to be processed in terms of a set of exemplars similar to it, $E' = \{e'_1, e'_2, \dots, e'_n\}$, e.g., *She sent me a photo*, *She emailed me a photo*, and so on. Here let the integration operation be a logical summation over a set, \bigvee , for convenience.¹⁾ Then the integration of E' , notated as $\zeta(E')$, can be represented as follows:

$$(17) \quad \zeta(E') = \bigvee_{i=1}^n e'_i = e'_1 \vee e'_2 \vee e'_3 \vee \dots \vee e'_n$$

In terms of this assumption EBCG marks a sharp contrast with the currently prevailing theory of construction in that in the latter the sentence is considered to be processed with abstract template called argument structure construction, specifically in this case, Ditransitive Construction. In this sense the prevailing version of Construction Grammar can be called *Abstraction-based Construction Grammar*, hence hereafter referred to as ABCG.

3.2 Background assumptions

As assumed in the exemplar theories, be it psychological or linguistic, EBCG also takes a *radically memory-based* view of human knowledge. Specifically, EBCG assumes that we humans memorize all the experience we have had as individual traces of memory. This assumption is sometimes called “rich memory” assumption (e.g., Port 2007a).

The motivation for adopting this assumption is that, as given in 2.4, in order to explain the process of learning, especially, how the frequency of linguistic input affects language learning, the rich memory assumption is almost necessary, because, without this, the theory needs additional constructs or hypotheses which explain how the learning begins, how the frequency counts, how errors are corrected afterwards, and so on (Cf. Arnon & Snider 2010:77; Bybee 2010:18).

3.3 Theoretical conceptions

3.3.1 Preliminary notes on an exemplar

In the previous chapter we overviewed some details of exemplar models mainly of categorical judgment. Those models represent an exemplar e as a multidimensional feature vector, which can be interpreted as a point in a fully parametrized multidimensional space. This way

¹⁾The definition and mechanism of meaning construction via integration of exemplars will be provided in detail in subsections from 3.3.8 to 3.3.10.

of representation is, however, considered to be largely due to the convenience of computation, given that those models are simulation models. In fact, non-simulation models such as *Episodic Grammar* (Borensztajn & Zuidema 2011) do not assume an exemplar as a multidimensional vector.

Conceptually, an exemplar *e* can be seen as an individual *trace* of (a fragment of) an experience in a memory including many kinds of information such as visual, auditory, emotional, and in some cases social properties. A little more specifically, it is possible to assume that *e* is an *individuated* or *tokenized* experience or, more simply, a *token of experience*.²⁾

Further, it can also be said that in order to individuate a token from a stream of experience, the token should be *recognized* as an *instance* of something previously experienced. In other words, tokenization would probably needs recognition of a *type*. In EBCG, therefore, exemplar is, rather roughly, defined as a *recognized token of experience*.

3.3.2 Linguistic exemplars

In linguistic contexts, an exemplar is often assumed to contain *symbolic* contents such as words and syntactic parses (e.g., Bod 2009). By symbolic it is meant that an experience stored in memory includes top-down information associated during the course of recognizing something. For instance when we perceive a stream of speech we not only experience auditory stimuli but also recognize what words are contained in the stream. The recognized words are, needless to say, things which were *acquired* or *learned* at some point in the past; they are couplings of bottom-up sensory data and top-down conceptual elements. In exemplar theory of language, therefore, both perceptual and symbolic contents involved in an experience are assumed to be stored together as an exemplar.

As many other exemplar theories of language, EBCG also assumes that symbolic information is included in an exemplar and stored in some way. More specifically, pieces of symbolic information are considered to be, as it were, *attached* to sensory data obtained during some experience, functioning as *indices* (Cf. Kuroda 2009:281). In consequence it is assumed that they can be *edited* or *reconstructed* afterward if necessary, in such cases as you had misheard an utterance and later you found it to be so or as you had interpreted an expression literally but later you noticed it was metaphorical.

²⁾Tokeniation is the term originally used in computer science, especially in the field of *Natural Language Processing* (NLP), referring to the operation of segmenting a text stream into any kinds of meaningful units such as words.

Here the concept of *linguistic exemplar* is introduced, which represents a kind of exemplar, that is, a recognized token of experience, including a recognition of some linguistic symbols. Linguistic exemplars, therefore, necessarily contain symbolic information. Note here that henceforth, unless otherwise specified, an exemplar is referred to as a linguistic one.

3.3.3 The unit of exemplar memory

At this point a question may arise: What is the unit of exemplar memory? How can we tokenize an exemplar from the continuous stream of experience? This question actually matters but, unfortunately, is yet to be given a precise answer by EBCG. There are, however, a few things which can be said as to the question about the memory unit.

If an exemplar is a recognized token of experience, the unit of it should have much to do with the process of recognition by humans. Recognition of things is, in turn, considered to be related with what can be called *attention* or *focus of consciousness*. If so, tokenization of an exemplar unit and our ability of attention are considered to be interrelated.

In fact, for example, Robert Nosofsky includes the mechanism of *selective attention* in his model of exemplar memory in order to explain the relation between classification and identification (Nosofsky 1984, 1986). In linguistic context, Wallace Chafe argues intensively about how important the focus of consciousness in human language processing, be it interpretation or production (Chafe 1994).

Further, Chafe (1994) makes a specific claim about the unit size on which consciousness is focused: “[a]pparently a focus of consciousness is typically expressed with *four* words of English” (Chafe 1994:65, emphasis added). This figure is obtained by calculating the average length, i.e., the number of words, contained in an *intonation unit*, which is regarded as a reflection of focus. Needless to say the specific figure, four, cannot be adopted as a typical length of linguistic exemplar only from Chafe’s (1994) argument, but it should be taken into consideration to lesser or greater degree.

Here, quite tentatively, a unit of exemplar is assumed to be an *linguistic experience of a sequence containing up to around seven words*. The figure *seven* comes from a famous argument about human *working memory* by George Miller, expressed as the “magical number seven plus, or minus two” (Miller 1956), which is considered as the limitation of the number of things we can pay attention to at a time.

3.3.4 A sentence as an episode

It should be said that the unit of this length is not guaranteed to correspond to such a linguistically meaningful unit as a sentence. In fact in many cases the length of a sentence exceeds the number, sometimes to a great extent. This does not mean, however, that linguistic units which have been long discussed in the history of studies of language are needless to assume. They probably have some kind of reality.

The problem here is, therefore, how to reconcile the assumption of unit length in EBCG with the traditional notion of linguistic unit like a sentence. It is well known that segmenting a speech stream into a sentence is quite hard to accomplish, which implies that sentences do not matter much for spoken language. In spoken language, instead, it is known that prosodic features such as intonation and meter do a similar job (e.g., Port 2003, 2007b), but, as mentioned above, prosodic units are considered to correlate with that of attention, hence mostly covered by EBCG's definition of the exemplar unit.

In written text, in contrast, we can often find *fullstops* or some other punctuations between a sentence, which clearly tells us the boundaries of the unit, but the length of a sentence is often much longer than that of exemplar unit. It is beneficial to assume that, therefore, in addition to an exemplar there is another kind of memory unit, which covers such a long unit as sentence. The unit is an *episode* (Cf. Borensztajn 2011; Borensztajn & Zuidema 2011). Episode is the term employed in psychology of memory, especially in the classification of memory into *semantic memory* and *episodic memory* (e.g., Tulving 1972).

Episodic memory is considered to have its *chronological order* of event in it, and the events contained by it are arranged with some overlaps (Eichenbaum et al. 1999). The episode and its internal events can be reinterpreted as a sentence and exemplars, respectively. Further, longer units than a sentence such as a paragraph and a text can be included in the framework of EBCG in terms of the concept of episode. Episode no longer is a memory trace; it can be seen as something mentally *reconstructed* based on the chronological order of stored experiences.

3.3.5 A brief note on words

Obviously a *word* is also assumed to do a job in exemplar tokenization. Here a word is considered as a unit of *segmentation*, which provides the whole unit with *multiplexity*. Multiplexity is the property which can be characterized as having component parts. Multiplexity is thought to play a crucial role in finding similarity between units, because similarity is in

many cases given as a kind of partial overlaps.

3.3.6 Construction effects: stored or not stored, that is NOT the question

As seen in 2.3.2, in the tradition of ABCG a construction has been defined as the stored unit of form-meaning pair. Exemplar-based view, however, does not concern whether a unit is stored in memory or not, because all the units once hear/read are assumed to be stored (Bybee 2010:24). As a corollary to this exemplar-based view of memory, EBCG declines the very assumption of construction as a stored unit in memory, because, if it is adopted, all the units memorized are to be considered as constructions, which obviously makes, at best, only a quite trivial sense.

Instead, a construction is considered only as a *phenomenon*: there do not exist any constructions, but only *construction effects*. The construction effect is a kind of *completion effect* in which given some input 1) the input is interpreted as some parts of a unit as a whole, 2) then the unit itself is associated or *evoked* 3) and hence some other *nonpresent* parts contained in the unit are completed. In the case of the example given in (16), for example, the discontinuous sequence [She ... me a photo] can be thought to evoke some whole such as *She sent me a photo* and therefore the part “sent” with the meaning accompanied with the part would be completed.³⁾

The unit assumed to be evoked is, of course, a stored exemplar. As seen in 3.3.1, an exemplar is defined as a recognized token of experience, and therefore it also functions as a unit for recognition. Further, a linguistic exemplar is assumed to be indexed with the interpretation of it as its meaning. Therefore, with exemplars not only can a novel input be recognized as a unit but also be provided with some interpretation, be it appropriate or not.

In this sense, an expression without any novelty can be, though it sounds somewhat strange, thought to be *completed*. For example, ABCG’s definition of a construction may allow such a phrase as *I love you* to be a kind of construction because the phrase is used quite frequently and therefore can be seen as stored in memory (Cf. Goldberg 2006:64). EBCG also assumes it as involved in construction effects because when we here someone say “I love you” we can easily recognize it as an instance of the phrase *I love you* which is, say, often used as an expression of one’s love. Here, importantly, such a recognition can only be achieved

³⁾Note that completion is not solely of forms nor of meanings, but of *units* including form and meaning at a time.

by completing some information missing in the input, i.e., the sound stream produced by someone.

3.3.7 Definition of the construction effect

However, only with a completion effect construction cannot be characterized or defined thoroughly. In fact a completion occurs, at least by definition, at every moment we humans recognize something as an instance of any object previously constructed in mind. The construction effect is just one type of completion effects.

In the first place, a construction effect is restricted to be the *linguistic* kind of completion, or, simply, a *linguistic completion*. Completions as in visual shape recognition, individual recognition and so on, are therefore excluded. Linguistic completions are defined as completions involving linguistic exemplars.

In the second place, linguistic completion of *simplex* units is also excluded. A unit evoked in a construction effect must be *multiplex*, that is, it must be, as the name shows, *constructed* from some component parts. The criterion of multiplexity is that some part of the unit can also function as a unit to be recognized. For example, words such as *such* and *as* can be considered as multiplex units, because their supposed component parts, that is, letters *s*, *u*, *c* and *h* or phonemes /s/, /ʌ/, and /tʃ/, are also targets of recognition. In contrast, the parts of them, in turn, cannot be seen as multiplex units, for their component parts are hard to assume. Needless to say, the sequence of the two words, i.e., *such as*, can be regarded as a multiplex unit.

Here the construction effect, notated as η , is defined as follows:

- (18) The construction effect η is a linguistic kind of completion effect involving an exemplar indexed with a multiplex linguistic unit.

The multiplex unit involved is called the *source of η* , represented as $\sigma_{i,j} \in \sigma_i$, and some parts associating the source and hence evoking the effect is called the *evoker of η* , ϵ_i . As described above $\sigma_{i,j}$ is an exemplar $e'_j \in E'$, that is, an exemplar set whose elements are similar to the given input s_k . Note here that it is not the case that all the exemplars in E' function as $\sigma_{i,j}$; only the subset of them do (see 3.3.9 for details).

Let a specific construction effect, as opposed to the construction effect in general, be represented as η_l . Then η_l can be introduced by the following triplet:

$$(19) \eta_l = \langle \sigma_i, \epsilon_i, s_k \rangle$$

Specifically, for a given input s_k , some parts of s_k function as ϵ_i and a subset of associated exemplars with them, E' , work as σ_i , resulting in a specific construction effect η_l (see Table 3.1).

Table 3.1: Components of η_l

	Examples
σ_i	$\{She\ sent\ me\ a\ photo,\ She\ emailed\ me\ a\ photo,\ etc.\}$
ϵ_i	[she ... me a photo]
s_k	“She skyped me a photo”

3.3.8 The two steps of meaning construction

As shown in 3.1, EBCG assumes that sentence processing is achieved by means of transferring the *constructed* meanings and/or forms based on exemplars associated with the given sentence. Here transferring simply denotes the process of, say, *recycling* past experiences in order to cope with something unknown, but, in contrast, meaning construction is a rather complex process which needs somewhat detailed explanations.

In (17) the integration of exemplars was represented as the logical summation for convenience, but it is only for convenience. In order to explain the process fully, it is needed to introduce a *two-step model* of meaning construction: the first step is what can be called the *summation* of σ_i , $\mathfrak{S}(\sigma_i)$, and the second, the *unification* of σ , $\mathfrak{U}(\sigma)$, where $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$, whose precise definition will be given below.

3.3.9 The summation of σ_i

The first step, $\mathfrak{S}(\sigma_i)$, is a kind of logical sum operation with a few additional assumptions. They are: 1) the *parallel association principle*, 2) the *non-contradictory condition*, and 3) the *largest set condition*.

Let us take an example:

$$(20) \text{ She kicked me a question.}$$

This sentence, though composed only of quite famous words, may sound somewhat novel and be a little hard to interpret simply. It is reasonable to assume that the sentence has two strong evokers of some η_l , that is, [she kicked me a ...] and [she ... me a question], represented as ϵ_1 and ϵ_2 , respectively. It is assumed here that other things being equal, the more words ϵ_i contains, the stronger it evokes some η .

The parallel association principle

The parallel association principle plays a role here. The two evokers are assumed to associate exemplars *parallelly*; that is to say, ϵ_1 and ϵ_2 are considered to associate different sets of exemplars *independently*. Let the associated set of exemplars be notated as σ_i^* where the index i corresponds to that attached to the evoker ϵ_i .

For example, ϵ_1 associates such exemplars as shown in (21), i.e., σ_1^* , and ϵ_2 associates those in (22), σ_2^* :

- (21) a. She kicked me a ball.
 b. She kicked me a lot.
 c. She kicked me a little.
 d. She kicked me a few times.

- (22) a. She asked me a question.
 b. She didn't ask me a question.
 c. She threw me a question.
 d. She gave me a question.

The non-contradictory condition

Then the two sets of associated exemplars are tried to be *summed* parallelly. The two conditions, i.e., the non-contradictory condition and the largest set condition, are applied here to the summation based on the parallelly associated exemplars. In essence, the former can be explained as follows: the summation operation only applied to any subset of σ_i^* every pair of whose elements are not contradictory with each other. The set with elements selected by this condition is σ_i . If the condition for any pair (x, y) of being not contradictory is represented as $P(x, y)$, σ_i can be defined as follows:

$$(23) \sigma_i := \{e \in \sigma_i^* \mid P(e_j, e_k)\}, \text{ where } 1 \leq j, k \leq |\sigma_i|$$

Note here that, in principle, in most cases σ_i^* provides more than one set of σ_i for a given ϵ_i . For instance, σ_2^* contains at least two sets whose members are not contradictory with one another. The two sets, conveniently notated as $\sigma_{2,1}$ and $\sigma_{2,2}$ respectively, are as follows: $\sigma_{2,1} = \{\textit{She asked me a question (= (22a))}, \textit{She threw me a question (= (22c))}, \textit{She gave me a question (= (22d))}, \dots\}$; $\sigma_{2,2} = \{\textit{She didn't ask me a question (= (22b))}, \dots\}$.⁴⁾

The largest set condition

This situation obviously causes a decision problem. The solution to the problem EBCG offers is the largest set condition. It is quite simple: if, for a given ϵ_i , there are more than one associated set whose elements are not contradictory with one another, the *largest set* can only survive. In the case of ϵ_2 , it is quite reasonable to assume that $\sigma_{2,1}$ is the largest, because there are few cases where the negative form of some verb occurs more frequently than positive one.

Due to the largest set condition, $\sigma_{2,1}$ finally obtains a status of σ_2 . The summated meaning of σ_2 can be described, rather roughly, as follows: a female person referred to as “she” provides or provided me with a question in such a manner as asking, giving, throwing, or something like that.

It should be noted that the summation operation $\mathfrak{S}(\sigma_i)$ is what brings about a construction effect, η_i . Put differently, η_i is evoked by summing a set of exemplars σ_i associated with from ϵ_i under the non-contradictory condition and the largest set condition.

The other set of exemplars associated with (20), namely σ_1 , also does its job. The exemplars contained in σ_1 are presumably not contradictory with one another, so it can be thought to be summed into one, resulting in such a meaning as follows: a female person referred to as “she” hit me with her foot, or transfer a ball or something to me, in a kicking manner. Simply put, therefore, $\mathfrak{S}(\sigma_1)$ only says that “she” did some “kicking” action on or toward me.

Now σ can be given a precise definition: $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ is a set of sets associated by a single input s_k each of whose member, σ_i , is a target of summation operation \mathfrak{S} , that is, a set associated by a source of construction effect ϵ_i contained in s_k whose members are not contradictory with one another.

⁴⁾It is also possible to assume that $\sigma_{2,2}$ contains such sentences as (22c) and (22d), because it is possible to say, for example, “She didn’t ask me a question; actually, she threw me a question.”

3.3.10 The unification of σ

Let us move on to the next step, $\mathcal{U}(\sigma)$. A set of selectively associated exemplar sets, σ , is assumed to be integrated via the unification operation \mathcal{U} . The unification here resembles one in computer science and mathematical logic, known for being adopted in the programming language `Prolog` as its fundamental operation. In the field of linguistics the formal theory of grammar known as *Head-driven Phrase Structure Grammar* (HPSG: e.g., ?) adopts the operation as its basic part of computation framework.

The unification $\mathcal{U}(\sigma)$ is defined as a process described below. For simplification the following explanation is in the case where σ has only two members, but it is not assumed that the operation only works pairwise:

- (24) a. **meaning decomposition:** for a pair of $\langle x = \mathfrak{S}(\sigma_i), y = \mathfrak{S}(\sigma_j) \rangle$, decompose each of the two into parts, resulting in an array $x^* = [x_1 x_2 \dots x_n]$ and $y^* = [y_1 y_2 \dots y_m]$, respectively;
- a'. the decomposition should be done so that the number of overlapping parts between those from the two, i.e., $|x^* \cap y^*|$, would become maximum;
- b. **proper alignment:** sort the order of the one whose size is smaller so that the order of its elements would maximally correspond to the other;
- c. **wildcard insertion:** for each array, insert a special element \bullet so that every element of the two would completely correspond;
- c'. the special element \bullet functions as a kind of *wildcard*, and therefore the logical product of something with \bullet always returns the thing itself, that is, $\forall x(x \wedge \bullet = x)$;
- d. **columnwise unification:** for each pair of i th element in x^* and y^* , that is, $\langle x_i, y_i \rangle$, make the logical product of the two, $z_i = x_i \wedge y_i$, and return the newly obtained array of the products, $z = [z_1 z_2 \dots z_l]$;
- e. sum up z and return the sum as $\mathcal{U}(\sigma)$.

The logical product $z_i = x_i \wedge y_i$ provides a *specification* of meaning to some *underspecified* part in x_i or y_i . By “underspecified” it is meant that the meaning is stated ambiguously or disjunctively. An underspecified part either in x_i or y_i is specified by means of a part of the other, if and only if the two parts are *not contradictory* with each other.

For example, in the case of $\mathfrak{S}(\sigma_1)$ and $\mathfrak{S}(\sigma_2)$ given above, notated as x and y respectively, the unification process can be described as follows:⁵⁾

⁵⁾The denotation brackets enclosing strings S with an index i , $\llbracket S \rrbracket_i$, is used here to represent the meaning of

(25) a. x is decomposed into:

[[she]₁ [provides or provided]₂ [(to) me]₃ [a question]₄ [in some manner]₅];

b. y is decomposed into:

[[she]₁ [hit or transferred]₂ [a ball or sth]₃ [(to) me]₄ [in a kicking manner]₅];

c. y is sorted into:

[[she]₁ [hit or transferred]₂ [(to) me]₃ [a ball or sth]₄ [in a kicking manner]₅];

d. for each corresponding pair a logical product is made:

[she]₁ \wedge [she]₁ = [she]₁, [provides or provided]₂ \wedge [hit or transferred]₂ = [transferred]₂, [(to) me]₃ \wedge [(to) me]₃ = [(to) me]₃, [a question]₄ \wedge [a ball or sth] = [a question]₄, and [in some manner]₅ \wedge [in a kicking manner]₅ = [in a kicking manner]₅;

e. sum up the logical products: [she transferred a question to me in a kicking manner]

Then the meaning obtained from (25e) is transferred to (20), resulting in an interpretation of (20). This may probably have a flavor of metaphor.

Here a new way of notation is introduced. The unification is conceptually similar to the *inner product* of two vectors, notated as $a \cdot b$ for given two vectors $\langle a, b \rangle$. $\mathcal{U}(\sigma)$ can therefore be represented as follows:

$$(26) \mathcal{U}(\sigma) = \mathfrak{S}(\sigma_1) \cdot \mathfrak{S}(\sigma_2) \cdot \dots \cdot \mathfrak{S}(\sigma_m)$$

As is evident from the very integration operation of σ , it is often the case that there are two or more types of construction effects involving in a single input. In (20) for example there are two construction effects evoked, which are roughly characterized as the Ditransitive Construction with a meaning of “asking” and the Ditransitive or Simple Transitive Construction with a verb “kick.”

It can be said that, therefore, $\mathcal{U}(\sigma)$ partly corresponds to the *fusion* discussed in Goldberg’s theory of construction (Goldberg 1995:50-52). The correspondence, however, only holds partly. As seen in 2.3.1, Goldberg (1995) assumes that the fusion is the process of integrating an argument structure construction and a verb. This means the process necessarily involves a verb, which is somewhat incompatible with the assumptions of EBCG because such abstract components as part-of-speech categories are declines under the *anti-abstractionist thesis*.

the strings.

3.3.11 When meaning construction fails

It is also assumed that there are cases where meaning construction fails. Let us take the famous example of a semantically anomalous sentence given by Chomsky (1957):

(27) Colorless green ideas sleep furiously. (Chomsky 1957:15)

EBCG is required to explain how the interpretation of this sentence fails.

First, clearly, in this sentence it is highly difficult to find any sources of construction effect containing more than one word. Therefore in (27) only the words contained in it can be regarded as sources of construction effects: $\epsilon_1 = [\text{colorless} \dots]$, $\epsilon_2 = [\dots \text{green} \dots]$, $\epsilon_3 = [\dots \text{ideas} \dots]$, $\epsilon_4 = [\dots \text{sleep} \dots]$, $\epsilon_5 = [\dots \text{furiously}]$.

Second, it follows that, therefore, the exemplars associated with the sources are almost all the sequences containing each of the words. This obviously leads to a situation in which each $\mathfrak{S}(\sigma_i)$, the summation of selectively associated exemplars, highly resembles what can be called the *lexical meaning*.

Third, at the phase of $\mathfrak{U}(\sigma)$, it is quite probable that the second step of it, *proper alignment*, does not work well. This certainly results in insertion of a large number of wildcards (28), and hence the summed meaning of an array of columnwise unifications makes almost no sense, resembling (29).

(28)	a.	[[[colorless]]	•	•	•	•]
	b.	[•	[[green]]	•	•	•]
	c.	[•	•	[[ideas]]	•	•]
	d.	[•	•	•	[[sleep]]	•]
	e.	[•	•	•	•	[[furiously]]]

(29) [[colorless]] + [[green]] + [[ideas]] + [[sleep]] + [[furiously]]

In summary, sentences such as (27) is analyzed as unable to interpret in terms of the meaningless output returned by the unification operation $\mathfrak{U}(\sigma)$. The meaninglessness is caused by the nature of proper alignment in the unification process, which requires the maximal correspondence of each parts in an array of decomposed meanings.

It should be noted that for many linguists, the sentence (27) sounds no longer anomalous because of the frequent exposures to the sentence. It is assumed that, therefore, the sentence is processed in whole, that is, for linguists the pattern [colorless green ideas sleep furiously]

functions as ϵ which associates a number of individual occurrences of the sentence, resulting in an interpretation like “this is a famous sentence which is logically or semantically anomalous, given by Chomsky in order to illustrate syntax and semantics are two totally different things.” In this case the meaning of it, rather trivial, can be represented as follows:

(30) $\llbracket \text{colorless green ideas sleep furiously} \rrbracket$

3.4 Methodology

In order for any theory to verify its assumptions and hypotheses, methodology is of great importance. Without a principled methodology any account for empirical data would fall into an ad-hoc one, causing it to be an armchair theory. As seen in Chapter 1, the aim of this dissertation is to achieve a theoretical investigation of EBCG, that is, to *construct* the theory, but this does not mean that EBCG has no concern about empirical research such as data analysis. Rather, it is assumed that the way of describing real data to verify the assumptions should be included in the theoretical components.

This section provides somewhat detailed description of EBCG’s way of analyzing data. It includes 1) the tool employed describing data, 2) the procedures of analyzing data, and 3) two *slogans* of analysis, “take the form-based approach” and “think larger, the better.”

3.4.1 The surface pattern

So far it has been shown that that EBCG is a theory of construction investigating construction effects, rather than constructions themselves. A specific construction effect η_l can be provided by the triplet $\langle \sigma_i, \epsilon_i, s_k \rangle$ as seen in (19). Therefore, it is obvious that EBCG’s concern is to explain some construction effect based on a given input s_k including its subparts functioning as ϵ_i and an associated set of exemplars $\sigma_i = \{\sigma_{i,1}, \sigma_{i,2}, \dots, \sigma_{i,m}\}$.

However, this way of explanation may sound, unfortunately, quite ad-hoc, because of its very dependence of on individual examples. In addition, there may be, and quite possibly actually are, more than one triplets which can give rise to the same effect η_l . For example, such examples as in (31) may also be involved in the same effect as that involved in the processing of (16) (= She skyped me a photo):

- (31) a. She iPhoned me a photo.
b. She Tweeted me a photo.

These two also contain a discontinuous part [she ... me a photo] and in consequence the part should function as ϵ which associates the same exemplars as associated in the case of (16), resulting in, of course, the same effect as evoked in (16).

This means that the construction effect, if defined in the way shown in (18), depends crucially on the *evoker* of it, namely, ϵ . For EBCG, therefore, ϵ should be the major tool for description. From this it follows that, methodologically, EBCG investigates ϵ s which effectively specify a construction effect η_l . When focusing on this methodological purpose, ϵ is alternatively called the *surface pattern*, notated as p .

3.4.2 Procedures

At least theoretically, it is better to be able to explain some general phenomena than to be able to only account for a number of specific examples. Therefore, EBCG, to begin with, aims at explaining major constructions in English discussed in many literatures.

In order to achieve this goal the following procedures are taken:

- (32) a. seeking one or more surface patterns which is assumed to be effective to identify a certain construction effect corresponding to one of the major constructions;
- b. based on the specified patterns, collecting actually employed tokens occurring in linguistic database called *corpus*;
- c. checking whether the collected data can actually be regarded as instances of the construction.

For example, the English *Ditransitive Construction* (e.g., *Freddy gave me the globe*) can be identified at least partially by a pattern [... will ... me a ...] as in *You will give me a kiss*.⁶ At the same time, however, tokens which cannot be considered as instances of this construction may also be found using the pattern, such as *You will miss me a lot*. If the number of those tokens are rather small, they are regarded as exceptions, but that is not all: then other factors which make them exceptional should be investigated. As for the “miss you” example above the pattern [... miss ... a lot] may play this role.

3.4.3 Take the form-based approach

Construction Grammar, exemplar-based or not, assumes that a construction has its own form and meaning. Therefore at least logically it can be said that there are two possible approaches

⁶This pattern can be obtained given the properties of typical instances (see Gries 2003 for details).

of Construction Grammar: form-based one and meaning-based one. The meaning-base approach is taken in a number of works by A. E. Goldberg (e.g., Goldberg 1995) as in explaining the nature of an argument structure construction.

In contrast, EBCG takes form-based approach, because it is assumed that the meaning of a sentence or the like cannot be given in the stream of speech or letters but be completed based on our memory of previous experience. This attitude comes from the *cognitive-realist thesis*, which in part requires the theory to be able to explain on-line language processing. Specifically, EBCG assumes that meaning-based approach can only be possible if language is seen as an *analyzer*, rather than a user of it, because language users cannot access meanings directly from the given input. If so, the analysis should also be form-based in terms of the cognitive-realist thesis.

3.4.4 Think the larger, the better

When describing linguistic data, it is possible to use as small as possible units for description, such as words, morphemes or in some cases phonemes. This attitude can be called a reductionist view of language, in which language is considered to be composed of rather small number of atomic elements and they are composed into larger one. In fact the majority of linguistic theories take this attitude.

EBCG, however, goes in just the opposite direction, assuming “the larger, the better.” This attitude is exemplified by the investigation of ϵ or a surface pattern p . As briefly noted in 3.3.9, it is assumed that other things being equal, the larger the pattern is, namely, the more words it contains, the stronger the pattern evokes some construction effect. In describing data, therefore, EBCG uses patterns containing as many words as possible.

This assumption comes from the notion of *specification* of meaning. In principle, any small fragments of a sentence such as words can evoke some kind of construction effect, but in many cases words are probably found to be weak to specify the meaning of sentence, as illustrated in 3.3.11. Using a web-search metaphor, in most cases the more words you use as a query of searching, the more specific and hence the more approximate to what you want the result becomes. It can be said that specification is also beneficial for language users, given the on-line nature of language processing.⁷⁾

⁷⁾Incidentally, this slogan shares much with the *maximalism* advocated by Ronald Langacker (e.g., Langacker 1987)

3.5 Advantages of EBCG over ABCG

EBCG has at least three advantages over ABCG: 1) it has less theoretical constructs than ABCG; 2) it can easily deal with the polysemy of a construction which enables us to avoid overgeneralization; 3) it enables us to discuss matters on representation and processing at one time, while ABCG does not. In the remaining of this section those advantages are discussed in some details one by one in this order.

3.5.1 Fewer assumptions

In ABCG constructions are assumed to be objects stored in mind. This assumption comes from the reasoning that without assuming the existence of constructions, many of linguistic phenomena such as systematicity of verb polysemy and possibility of novel sentence interpretation could not be explained.

It should be point out, however, that this way of reasoning is based on the fallacy of confusing *sufficiency* with *necessity*. It is true that assuming constructions as the stored objects, many phenomena *can* be explained, but it is not always *necessary*. There can be other possible assumptions which can equally explain the same set of phenomena.

Note here that, wherever there is a construction there also is a construction effect. There is a principle in philosophy of science called *Occam's razor*, which is the principle of fewest assumptions: other things being equal, hypothesis with fewer assumptions are better. Considering this principle, it follows that a theory without assuming constructions as objects is as better. Hence EBCG can be said as better than ABCG in terms of the number of assumptions.

3.5.2 Avoidance of overgeneralization

There is a situation in which a single argument structure construction has two or more similar but distinctive senses. This is called *constructional polysemy*, discussed in ABCG (e.g., Goldberg 1995:32–39). For example, Goldberg (1995:75) provides six senses in Ditransitive Construction and explains how they are related:

- (33) a. 'X CAUSES Y TO RECEIVE Z' (e.g., Joe gave Sally the ball.)
 b. Conditions of satisfaction imply 'X CAUSES Y TO RECEIVE Z' (e.g., Joe promised Bob a car.)
 c. 'X ENABLES Y TO RECEIVE Z' (e.g., Joe permitted Chris an apple.)

- d. ‘X CAUSES Y NOT TO RECEIVE Z’ (e.g., Joe refused Bob a cookie.)
- e. ‘X INTENDS TO CAUSE Y TO RECEIVE Z’ (e.g., Joe baked Bob a cake.)
- f. ‘X ACTS TO CAUSE Y TO RECEIVE Z at some future point in time’ (e.g., Joe bequeathed Bob a fortune.)

However, it can be pointed out that the constructional polysemy is caused possibly by a kind of overgeneralization, that is, the assumption of the argument structure construction as an object. In contrast, if the exemplar-based view is taken there is no need to discuss the polysemy, because the argument structure construction is considered as phenomenal and its meaning is assumed to come from associated exemplars.⁸⁾

3.5.3 Integration of representational and processing matters

ABCG assumes that a sentence is processed by fusing an argument structure construction with a verb, as explained above. However, it has to be said that under ABCG’s assumption the *recognition* of construction can hardly be explained, without appealing some kind of external devices. How can we recognize a given input sentence as an instance of such and such an argument structure construction?

In contrast, EBCG can easily explain how a sentence, especially a novel one, is processed. Rather, it may be better to say that EBCG is a model both of representation and processing in the first place, while ABCG is that only of the former. This obviously is a strongly advantageous point in EBCG.

3.6 Scope of the study

EBCG is a theory of grammatical construction and therefore, as manifested in Chapter 1, takes the constructionist thesis. This automatically leads to a relative disinterest in the internal structure of a sentence, which needs analysis, or decomposition, of a sentence into its components. As discussed above in 3.4.3, EBCG focuses on the formal, as opposed to semantic, aspect of language. Semantics is of secondary interest in EBCG, in the sense that it is dealt with only in an indirect way. Furthermore, somewhat related to the treatment of semantic aspects, EBCG does not address the issue of *acceptability* of a sentence, unlike the large ma-

⁸⁾As for this topic, see Yoshikawa (2010) for detailed discussions.

jority of grammatical theories. This section describes how internal structure of a sentence, semantics and acceptability are treated in the framework of EBCG.

3.6.1 On internal structure of a sentence

As repeatedly mentioned, the direct objectives of EBCG is the types of expressions. The internal structure of a sentence is, therefore, not pursued in an explicit and systematic way. Internals are considered to be provided by means of *element alignment* over segmented elements, as discussed in Radical Construction Grammar (RCG) (e.g., Croft 2005). The segmentation is achieved both on formal and semantic sides, but semantic aspects is prior in the sense that interrelations between semantic elements such as an actor of an event described by a sentence attribute interrelations between formal segments, resulting in syntactic properties, or *syntactic role* in the term of RCG, such as Subject. Figure 3.6.1 shows RCG's view of syntactic structures. The reason why the syntactic relation is represented as a dashed arrow is that RCG does not assume syntactic relations *per se* as mentioned in 2.3.1.

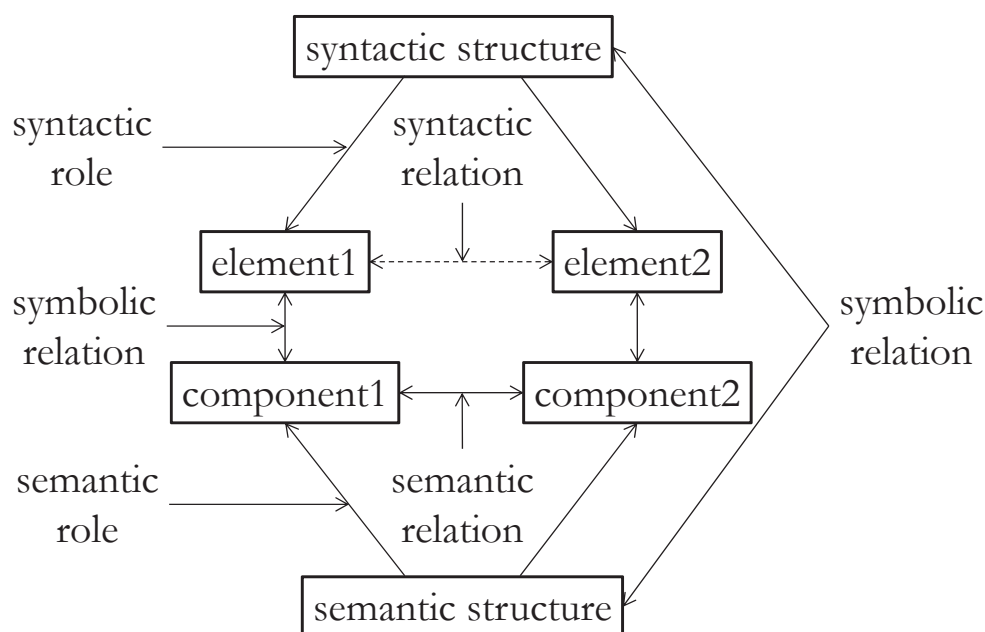


Figure 3.1: Alignment of syntactic elements based on semantic relations
(based on Croft 2001:176, Figure 5.1)

To take as an example a highly conventionalized English expression, *I love you*, RCG may explain that the phrase is first learned as fixed one without any component parts both in terms of its form and meaning but soon after the semantic structure of it is analyzed as containing three parts, namely, a person who loves, another person who is loved and affection felt by

one person for another, those parts are made to correspond to formal segments previously obtained in some way, *I*, *love*, and *you*.

The problem is, obviously, how to appropriately align semantic components with formal elements. A related topic are discussed in 3.3.10, 3.6.4 and 4.4.2, the latter two of which deal with the concept of *unification*, but the alignment problem itself is not directly dealt with.

3.6.2 On semantics

Semantics in EBCG is not a target of analysis. This does not mean, however, that EBCG does not concern semantic aspects at all. Semantics is, on the contrary, seen as important in construction processing, especially of understanding. Simply, semantics is *presupposed*, namely, regarded as something implicit but given. Every exemplar should be accompanied both with formal and semantic properties if language is involved. The semantics in EBCG is, therefore, this semantic properties attached to exemplars. Here a brief note is given on the nature of “semantic properties” in exemplar memory.

Intended and incidental associations

Exemplars assumed in exemplar theory are accompanied with rich information ranging from sensory data to category labels, as repeatedly mentioned above. Suppose an exemplar e_1 is accompanied with information including a linguistic sound sequence representing a furry animal, [kæt], and, simultaneously, a visual image of the animal, as well as an environmental noise caused by wind and trees, a situational label of the place (e.g., “at the park near home”), a label of the person who produces the sound (e.g., “mother”), and so on. All of the accompanied pieces of information are connected to e_1 with associative links and stored in memory. They can be, at least theoretically, re-associated with another exemplar, e_2 , if it shares some property attached to e_1 , such as the sound sequence [kæt].

Of what kind of experience is this exemplar? This may be an encounter of a scene in which a female happens to find a cat in a park and says “A cat!”, pointing the cat with her finger. Children before the age of 9 months may be unable to understand what the female does, at least in the same way in which adults do. To children at that stage the linguistic sound sequence [kæt] does not function as a formal aspect of a symbol, that is, the pronunciation of the word *cat*, but only is one piece of information associated with the exemplars. To adults or children older than one year, however, the sound sequence probably has a special status.

It automatically associates, or *evokes*, past experiences of seeing, thinking about, and even touching cats, even if those experiences do not involve the same sound sequence, as well as, say, an *image* of the animal characterizing it as a small, cute, and solitary being.

This means that the sound sequence is directly connected with a generalized notion connected to the experiences related to the animal, which is usually called a *concept*. This special kind of connection between a sound pattern and a concept is a *symbolic relation*. As will be discussed in 3.6.4, the acquisition of symbolic relations needs the understanding of others' intentions (Tomasello 2003:23).⁹⁾ In other words, if the association between a certain sound pattern with a certain concept is found to be *intended* by others, the association becomes symbolic. We may therefore call that type of association as *intended associations*, as opposed to *incidental associations*. Important here is that this characterization of symbols presupposes the existence of a concept paired with a sound sequence. Without preestablished concepts any sound sequence cannot be formal sides of symbols. Then the next question is: how can we build concepts? This question is equivalent to ask what the process of acquiring semantics is like. A tentative answer is provided just below.

Acquisition of semantics

In order to answer this question, we should first clarify what semantics is, or at least, what semantics is assumed to be in EBCG. Simply put, EBCG assumes that semantics is *the system of intendedly associated concepts via linguistic forms*. There actually are concepts which exist independently of any linguistic information. In fact even prelinguistic children can construct concepts on their own (reviewed in Tomasello 2003:20, also mentioned in 3.6.4), so semantics cannot be identical to the system of concepts in general. Instead, the components of semantics are considered to be special kind of concepts, which are obtained from top-down inference based on others' intentions accompanied with linguistic forms of any type.

In this sense semantics becomes available only after children begins to understand, or *discover*, others' intentions. Once they discover that there is such a thing as an intention, they become able to infer what is intended when others do something. This top-down inference then forms a solid relation between a concept and a formal pattern. This is also related to the process of *attention*, as described in 2.2.1 especially in the description of Nosofsky's (1986) model of categorization and identification.

⁹⁾He remarks that "[s]ounds become language for young children when and only when they understand that the adult is making that sound with the intention that they attend to something."

3.6.3 On acceptability

As described in detail in 3.4.3, EBCG concerns the mechanism of *meaning construction*, namely, the process in which a language user constructs a meaning from the confronting formal (and contextual) input. From this it follows that it does not put much focus on any kind of *meta analyses* or *evaluations*, represented by acceptability judgment, by language users on an expression such as a sentence; those are considered to be done based on the constructed meaning and hence after the construction. This means acceptability judgment can be seen as relatively peripheral, secondary thing compared to meaning construction.

First let us compare the following sentences (Chomsky 1957:15):

- (34) a. The book seems interesting.
 b. *The child seems sleeping.

The two sentences look similar, but only the former is judged as acceptable or grammatical. The latter is judged as ungrammatical probably because we know the verb *seem* cannot take a present participle as its complement. This way of explanation of the grammaticality judgment can be called the *grammar-based* explanation, but the judgment can also be explained in an *exemplar-based* way.

In the latter sentence we find the sequence [the child seems ...] which itself is not in any sense ungrammatical as seen in, for example, *The child seems hungry*, and the partial sequence would automatically make us think of a number of previously encountered sentences also the sequence, such as *The child seems hungry*, *The child seems happy*, *The child seems to enjoy*, and so on. At the same time, we also find another sequence in the sentence, namely [the child ... sleeping], and this probably associates us with such sentences as *The child is sleeping* and *The child was sleeping*. Based on the two sets of associated exemplars, we can reach a proper semantic representation of the sentence. Then, importantly, based on the obtained semantic representation we can make an inference in the following way: in order to express a meaning like this, we have to say otherwise, like “The child seems sleepy” “The child seems to sleep.”

This *possibility to say otherwise* is considered to be the factor for us to judge a certain sentence as ungrammatical. Put differently, it is assumed that even when we hear a sentence, we can *simulate* a representative *speaker* of the sentence and make inference on the natural way of expressing the intended meaning by saying the sentence. This is for now only speculative, but is a current explanation for the judge of a sentence as unacceptable or ungrammatical, named the *speaker simulation hypothesis*.

3.6.4 On the process of acquisition

As for the process of learning exemplars, EBCG basically adopts the idea of memory consolidation discussed by Borensztajn (2011), so at least in this dissertation the process itself is not a topic for discussion, but only an underlying assumption. Here, however, a brief description of the acquisition scenario EBCG assumes is presented, although some parts of it are only speculative argumentations.

Sound pattern encoding and concept building

At the very beginning of language acquisition, children do not know what linguistic sound is. However, due to their innate built-in ability, they can perceive and store auditory information of linguistic sound distinctively, in the sense that that type of sound is recognized differently from other types (e.g., natural sounds), even though they do not know the difference between linguistic and non-linguistic. The sounds are probably encoded in a high dimensional way, resembling a spectrogram (see Fig. 3.6.4).

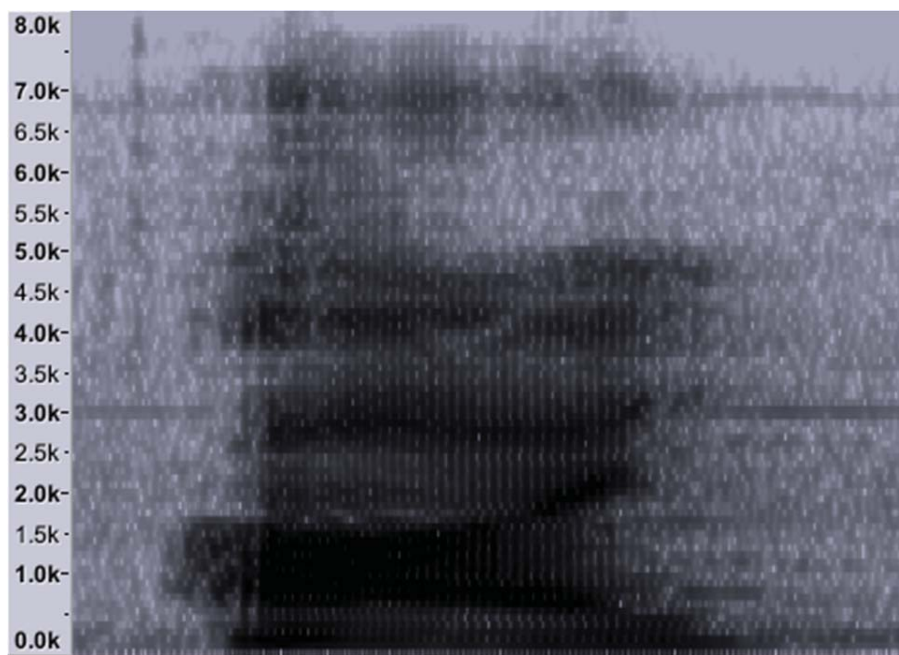


Figure 3.2: A sample spectrogram encoding a voice saying “Hi.”
(created using Audacity¹⁰⁾ [version 2.0.5])

In this period, probably during the first a few months children cannot group the individual sounds together in a phonological way. The sound information attached to exemplars are

¹⁰⁾<http://audacity.sourceforge.net/>

perhaps clustered with concrete auditory properties available from the birth or at least at very early stage of life. In fact a number of experimental studies mainly in the field of *Artificial Language Learning* (e.g., Saffran et al. 1996a,b) show that children in the prelanguage stage can find abstract sound patterns. From those findings, it is estimated that sound patterns can be used at the age of seven months at the latest (Marcus et al. 1999). What exactly is meant by the “sound pattern” here is not clear, but, at least on the behavioral level, we can assume children at round that age utilize that kind of information in whatever way.

It can be assume that there is another set of information sources available to young language learners: hypothesized labels the children themselves invent and attach to exemplars. The labels can function as category markers by integrating scattered exemplars into one group, not necessarily with shared perceptual properties. A children can hypothesized a label if there are some other type of properties shared by a group of exemplars such as the information on the place where utterances are produced and the person who produces the utterances, if the children has already acquired such concepts. In short, children utilize any kind of information available from innate or pre-acquired knowledge to bundle scattered exemplars together in this period, resulting in sporadic and idiosyncratic formations of category-like clusters.

During the same period, children also begin to acquire some conceptual knowledge. As Tomasello (2003:20) briefly reviews, it is found that children at the age of as early as 4 or 5 months already “have [...] formed concepts of simple objects and events.” At this age, however, they cannot produce and understand language, even though it is already possible for them to associate two aspects of one event such as auditory and visual information accompanied with the event. It is assumed that children at this stage only have associative relations between two different types of information, obtained via connections to exemplars. The left diagram of Figure 3.6.4 depicts how semantic and auditory patterns are linked associatively in exemplar memory.

Symbolic encoding

At around the first birthday, children eventually begin to understand and produce language. Just before this period, it is reported that children start showing signs of understanding symbolic relations, arbitrary connections between two essentially different pieces of information, as, for example, seen in the use of *symbolic gestures* (Tomasello 2003:34). This means that at this period children become able to learn conventional associations between a certain sound

pattern and a certain concept. As Tomasello (2003) argues, learning linguistic symbols needs socio-cognitive abilities such as intention-reading, which enable children to know what aspects or entities are worth encoding for the community the children are embedded in.

EBCG reinterprets this process of symbolic learning in the following way: children gradually accumulate hypothesized categories on their own and gradually find that some of them are also employed by other members as well. If the categories found to be shared by others go across two aspects of information in arbitrary relationship, they become symbols. What is important here is to learn symbol needs to know, or more precisely, come to believe in, something socially-shared. Without such sharedness, there is no necessity to make symbolic connections between two distinctive types of information.

It should be mentioned that the idea of viewing word learning as a process of hypothesis-testing is denied based on some facts on cognitive development (summarized in Markman 1990:57-58). This, however, does not lead to the denial of the assumption described above. What is denied in word learning literatures is the possibility of *inductive learning* based on hypothesis testing, not the ability to form hypotheses. EBCG assumes a process of *abduction*, not of induction, as a basic strategy children use in acquiring linguistic symbols (Cf. Roberts 2004). Hypotheses are not considered to be formed not inductively from actual experiences directly related to the target to be learned, that is, a word; they can be made, as argued above, based on some related but different types of information from the direct target such as any correlating situational features (e.g., the place where the word is produced and a person who produces the word). Once formed, the hypotheses can function as sources of inference in just the opposite direction to the case of induction: the inference is made because of the existence of the hypotheses.

Sub-symbolic encoding

The fact that children can utilize sound patterns at prelanguage stage does not mean that *phonology* is acquired by children at that stage. It is pointed out that we can hardly define a phoneme in a truly acoustic way because there is no invariant auditory properties across all the members of one single phoneme (e.g., Port 2007a, 2010). What actually happens would probably be the process of top-down categorization based on hypothesized labels. The labels are seen to be available via concepts- and situations-based grouping, by means of finding what is differentiated by other members of a community to which the children belong. Also important here is that the differentiation should necessarily be global. For example, if a child

can distinguish /d/ in *do* (/du:/) from /t/ in *to* (/tu:/), the child does not necessarily identify the /d/ sounds in *do* and *dee* (/di:/) (Port 2007a:151).

This goes along with an idea of phonology called *whole-word approach* and a theory called *Templatic Phonology* (e.g., Vihman & Keren-Portnoy 2013). The basic idea of the whole-word approach is that “children never learn sounds: they only learn words, and the sounds are learned through words” (? :148). Words are obviously symbols, so the learning of phonology should follow, as opposed to precede, the learning of linguistic symbols. This may radically reverse the relations between the two levels of articulations in a *double articulation* view, at least from the perspective of language learning. A schematic representation of word-based phonology formations is illustrated by the left panel of Figure 3.6.4. Auditory aspects of symbols which partly overlap with each other are, if effectively differentiating the difference of semantic aspects of them, sources of phonological characterization.

Morphology can also be characterized in a similar fashion, but in the case of morphology partial overlaps should also involve the semantic aspects of symbols. As Bybee (1995:429) remarks, morphology is characterized by “[p]arallel sets of phonological and semantic connections, if they are repeated across multiple sets of words,” as depicted in Figure 8. Morphology emerges where semantics and phonology of symbols *resonate* with each other.

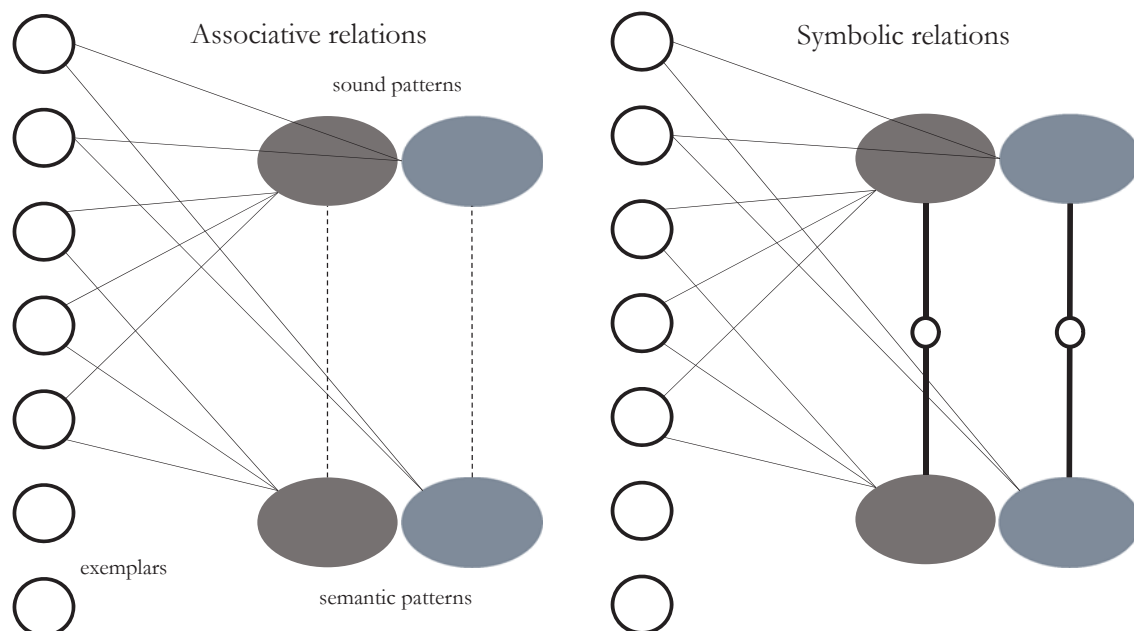


Figure 3.3: Associative relations and symbolic relations in exemplar memory

Morphology can be characterized as a *meta-symbolic association*, in the sense that morpho-

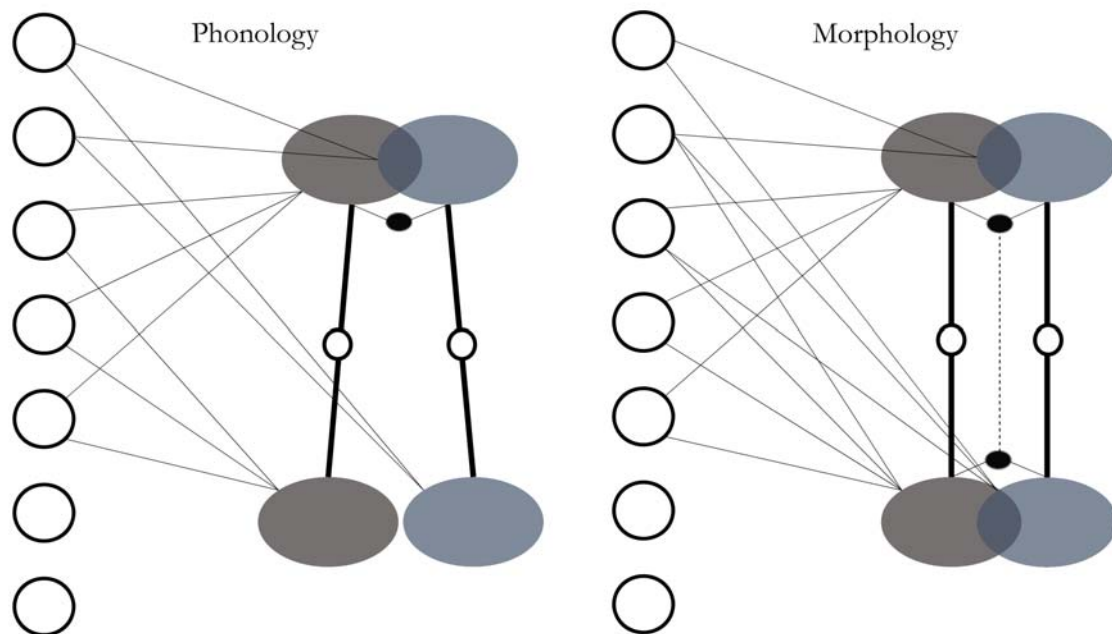


Figure 3.4: Phonology and morphology in exemplar memory

logical system represents a form-meaning association derived from two or more symbols. The same can be said in terms of *syntax*, an organization of

Tomasello (2003:125) characterizes syntactic structures as *second-order symbols*, in the sense that they are symbols of symbols. In English, the word order of [Subject Verb Object] typically denotes a meaning of the subject in some way causes effects on the object and this is shared by all the English speakers. Without this shared knowledge, no one can distinguish between *John hit Mary* and *Mary hit John*

3.7 Concluding remarks

3.7.1 Brief summary

This chapter presents the theory of construction called *Exemplar-based Construction Grammar* (EBCG) in detail. EBCG is a theory which assumes that 1) we human memorize all the exemplars of sentences previously heard/read; 2) when processing an input sentence we *associate* those concrete exemplars with the input; 3) then we *construct* meanings (and perhaps, forms) via *integrating* or *blending* the associated exemplars and *transfer* the integrated meaning (and form) to the input.

An exemplar *e* is defined as a *recognized token of experience* including symbolic informa-

tion. Especially, EBCG concerns the linguistic kind of exemplar, namely, the *linguistic exemplar*, which is considered to be a *word sequence* containing up to around seven words.

In EBCG the concept *construction* is reinterpreted as a phenomenon called the *construction effect*, η , which is a kind of completion effect evoked by some continuous or discontinuous patterns found in an input sentence, s_k . What evokes a specific construction effect, η_l , is called the *evoker* of η_l , notated as ϵ_i . This associates a set of exemplars and they are selected under a few conditions. Of the selected exemplar sets the largest one is called the *source* of construction effect, σ_i . Therefore, a construction η_l can be represented as a triplet $\langle \epsilon_i, \sigma_i, s_k \rangle$.

The meaning of a sentence is assumed to be provided by *integrating* the associated exemplars by $\epsilon = \{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$ obtained from an input s_k , that is, $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_m\}$. Specifically, first, for each i σ_i is *summated* into one with a kind of logical summation operation; secondly, the summated meanings are *unified* into one, resulting in an interpretation of s_k .

3.7.2 Future issues

Note that the model currently focuses mainly on the process of *comprehension*, as opposed to *production*, of a sentence. Therefore, it is pointed out that solely with the proposed model the production process of a sentence may not be explained in a principled way. Possible solutions to this may be either or both of the following two: to assume the production process as the reverse of the comprehension process or to attribute the mechanism of the production to some formal aspects of exemplar which are different from those utilized in comprehension.

If the former is true, production is considered to be done in the way in which, first, we think of semantic contents of what we would like to say or write, then, second, with the contents the stored exemplars are associated, and, third, the associated exemplars are integrated into one, resulting in an utterance to be produced. The problem here is that when producing linguistic expressions, we have to take into account the *linearity of language*. Comprehension does not require us to do so because the target of comprehension, the meaning, is not something aligned in a linear order. The target of production, in contrast, requires us to align elements in a linear order, which itself has in many cases a significant meaning. This problem can be rephrased as the problem of *syntax*, because the linear alignment of linguistic elements according to the meaning the speaker/writer would like to express is exactly the matter on syntax.

The problem of syntax can probably be solved if the latter is assumed as a process of

production. Suppose you would like to say something in response to someone who just said to you, "You look good." You may think of utterances based on the semantic, communicative, and some visual and auditory information, but you can also utilize yet another information source, namely the *linguistic context*. The minimum linguistic context available in this case is the utterance produced you were told, "You look fine." Only based on this, you are probably able to access a number of stored exemplars which were once produced by someone, perhaps including you yourself, just after the exact expression. As an approximation, the corpus search of [you look good .] on COCA tells us that the sequence is followed by such utterances as "Thank you so much," "You look so good," "Really?," "Yeah" and so on, when the speakers are changed. If such utterance sequences are stored in mind in some way, this kind of linguistic-context-based formal approach to production is considered attainable.

However, the latter approach also has a problem. It is assumed that there are many cases where the linguistic context does not associate us with any reasonable stored exemplars. The cases include those in which the context is a novel one. Therefore it would be appropriate to reconcile the two approaches into one. If exemplar theory is correct, any formal contexts available would be utilized to process anything, be it a linguistic expression to be comprehended or a semantic content to be expressed in a linguistic form. It is, therefore, highly natural to assume that both of the two approaches examined here are taken.

Chapter 4

Exemplar-based accounts for English Argument Structure Constructions

Constructions are the interface between language as experienced and language as represented in the brains of its speakers.

Taylor 2012:9

This chapter provides some detailed analyses of major grammatical constructions in English based on the theory and methodology introduced in the previous chapter. The constructions are: the *ditransitive construction*, the *resultative construction*, the *caused-motion construction*, and the *way* construction. Examples of those constructions are shown below:

- | | |
|--|------------------------------|
| (35) a. Freddy gave me the globe. | (ditransitive construction) |
| b. Mark pushed the door open. | (resultative construction) |
| c. Remy threw the book into the water. | (caused-motion construction) |
| d. Roy made his way through the crowd. | (<i>way</i> construction) |

They are called *Argument Structure Constructions* (Goldberg 1995:5). Argument structure is a typology of a predicate in terms of the number and the types of its arguments. Arguments can be seen as what fill the slots the predicate provides. For example, the English verb *give* is generally considered to provide three slots: *Subject*, *Object*₁ and *Object*₂; therefore the verb has an argument structure ⟨Subject, Object₁, Object₂⟩.

4.1 Introductory notes

It is reasonable to assume that *Argument Structure Construction* (ASC) is highly important both to *describing* English grammar from the *analyzer* point of view, and to *using* it from the

user point of view. For the former, ASCs provide a tool to organize an endless variety of sentences in a principled way, without appealing to the internal structure of a sentence. As for the latter, language users can utilize ASCs in order to predict the meaning of any unknown sentence, also in this case without parsing each sentence. In view of the creativity of language, which causes the endless variety of sentences, the latter aspect is of greater importance, because language users are almost always confronted with the meaning-prediction problem which should be solved in an extremely short period of time.

In consequence ASCs are worth analyzing under the *cognitive-realist* thesis. However, as argued in the previous two chapters, the current theory, *Exemplar-based Construction Grammar* (EBCG), does not assume the existence of any abstract constructions including ASCs. Instead, ASCs are simply reinterpreted the *major types of sentences* which need to be categorically judged. In other words, the objective of EBCG in terms of ASCs are only the problem of *categorical judgment* of major types of sentences.

With this in mind, some introductory remarks are provided in the remainder of this section below. Specifically, first, the methodology taken in the analyses presented in this chapter is described, and then disclaimers are added.

4.1.1 Methodology

Surface patterns

As provided in the previous chapter (3.4), the methodology taken by EBCG is largely based on corpus research in which one or more specified *surface patterns* are investigated. Surface patterns are specified based both on some pilot researches and on the findings from previous studies. For example, the surface patterns of the English ditransitive construction, discussed in 4.2, are specified based on previous studies on argument structure in general (Du Bois 2003) (see 4.2.1). On the other hand, the English caused-motion construction, discussed in 4.4, is analyzed based on surface patterns heuristically obtained through pilot researches. More precisely, the analysis of the caused-motion construction construction can be almost identified with the heuristics itself (see 4.4.1). In this way the method to find surface patterns is not established in an algorithmic manner, that is, a rigorous, rule-based fashion, but is fundamentally heuristic.

Statistics

There are some studies on grammatical constructions which focus on statistical indices computed from frequency information. The representatives of this approach are those by Stephan Th. Gries and Anatol Stefanowitsch, so called *Collostructional Analysis* (e.g., Stefanowitsch & Gries 2003; Gries 2003; Stefanowitsch & Gries 2005; Stefanowitsch 2008). They advocate the empirical thesis of language learning and representation taken under the usage-based theory of language, and assume that statistics or statistical analyses can provide a solid foundation of the theory (e.g., Gries et al. 2005:665-666). In this sense their approach is similar to the current theory.

However, EBCG does not take such a statistics-based approach. The reason for this is two-fold. First, it can be pointed out that there is no consensus about what kind of statistic index is suitable to analyze linguistic phenomena related to grammatical constructions. In fact, the validity of the statistic index employed in Collostructional Analysis is doubted by some scholars (e.g., Bybee 2010:97-101). Therefore it is at least risky to base analyses thoroughly on statistic properties of any kind. Nonetheless, this does not mean the rejection of any statistical analysis. EBCG utilizes statistics when it is useful and effective.

The other reason is more theoretical. Namely, the very nature of exemplar-based-ness may be somewhat incompatible with statistics-based approaches. At least in terms of the learning aspect of constructions, statistics should be treated in a careful manner, as seen in 2.3.3. It is impossible for EBCG to argue, for example, that the existence of a certain constructional effect is proved or supported by the high score in some statistical index, because that sort of argument is incompatible with the rich memory assumption of exemplar theory. Under the rich memory assumption, exemplars even with only a tiny statistical value should be considered as usable if condition is right. In other words, exemplar models do not work in a statistic way, though they may contain some statistical computations (e.g., Borensztajn 2011; Borensztajn & Zuidema 2011).

4.1.2 Disclaimers

Before entering into the analysis of constructions, some disclaimers are provided here. They are on acceptability and on the meaning of a construction. The former is a kind of manifestation that EBCG does not pursue any boundary conditions or rules which clearly exclude all and only the unacceptable sentences, as discussed in 3.6.3; it only concerns the mechanism of

meaning construction, as intensively explained in the previous chapter, especially from 3.3.8 to 3.3.11. The latter is, somewhat related to the former, an attitude in which analyses are performed based on the *form*, as opposed to the *meaning*, of a construction and therefore the meaning of a construction is not given much attention to; this is also related to some arguments about methodology of EBCG provided in 3.4.3.

On acceptability

As noted in 3.6.3, EBCG does not concern matters on acceptability; more precisely, EBCG does not provide any explanation about the acceptability of a sentence. It does not mean that acceptability is seen as trivial and hence there is no need to explain it; EBCG indeed assumes that acceptability is a highly important factor of language. However, as opposed to the majority of grammatical theories, EBCG considers that acceptability comes from factors external to grammar itself such as *speaker simulation* (3.6.3). This stance leads to an avoidance of attributing acceptability to any kind of grammatical properties. It can be said that EBCG's explanation provides a *precondition* on which acceptability is judged. In order for us as a user of language, not as an analyzer of it, to judge acceptability of a sentence we need some evidence on which we make a judgment.

On the meaning of a construction

In previous chapters (3.6.2 and 3.4.3) it is clearly stated that EBCG does not directly deal with the semantic aspect of constructions, both in terms both of theoretical and methodological perspectives. If some arguments about semantics of constructions are provided in EBCG's framework, they are, as it were, about *external* semantics, namely the interrelationship between constructions in terms of semantics, not about *internal*, or conceptual, semantics. For example the problem of *constructional polysemy* on the English ditransitive construction is argued in 4.2.2, but what exactly each sense of the construction is like is not discussed.

4.2 The ditransitive construction

The first construction to be investigated is the *ditransitive* construction, also known as the *double object* construction. Examples are the following:

- (36) a. She gave him a hug.
b. He offered me a job.

- c. I asked you a question.
- d. You taught us a lesson.

The structure of it is usually represented as [Subject Verb Object₁ Object₂] (Cf. Goldberg 1995). The typical examples are those including the verb *give* like (36a). The meaning of the construction is often explained as follows: the subject referent *X* (e.g., “She” in 36a) causes the referent of object₁, *Y* (typically a human), to receive something, represented as object₂. This is symbolically described as [X CAUSES Y TO RECEIVE Z] by Goldberg (1995:3). The construction has long been an object of study in the field of grammar theory. This is partly because of the uniqueness of the argument structure it has: It is the only construction in English which has two objects.

One of the representative studies on the construction concerning some cognitive aspects is Gropen et al. (1989), which investigates how children generalize and learn the constraints on the construction. They focus on the phenomenon called *dative alternation*, in which the two constructions, the *(to) dative construction* and the ditransitive, assumed to encode one and the same sense participate and behave as if they alternate with each other. The constraints are mainly related to the semantics of verbs. Sentences presented in (37) and (38) illustrate what dative alternation is like (Gropen et al. 1989:204):

- (37) a. John gave a gift to Mary.
 Hal told a story to Sue.
- b. John gave Mary a gift.
 Hal told Sue a story.
- (38) a. John whispered a secret to Sue.
 I donated a book to the library.
- b. *John whispered Sue a secret.
 *I donated the library a book.

Those in (37a) and (38a) are examples of the dative construction, and the others are ditransitive. As contrastively shown, verbs in (37) can participate in this alternation, but those in (38) cannot. Also related is the semantic type of Object₁ (Bresnan 1978, cited in Gropen et al. 1989:207):

- (39) a. I sent a package to the boarder/the border.

- b. I sent the boarder/*the border a package.

This contrast suggests that the recipient in the ditransitive construction should be a *person*, and in fact it is found that the construction encodes the meaning describable as “*cause to have*” (Gropen et al. 1989:203)

Other representative studies on the construction from cognitive perspectives are those by Newman (1996) and Mukherjee (2005). The former argues intensively about the semantics of the construction based on the theoretical framework called *Cognitive Grammar* (e.g., Langacker 1987, 1991). His characterization of the construction is based on four cognitive/semantic domains, namely 1) the spacio-temporal domain, 2) the control domain, 3) the force-dynamics domain, and 4) the domain of human interest (Newman 1996:Chapter 2). The four domains consist of a *complex matrix* in the sense that any parts, not necessarily all, of them are present in the semantics of the construction. For example, (40a) can be analyzed as involving all the four domains, but (40b) cannot, because the latter sentence describes a scene which does not include spatial transfer of things and any force-dynamics, only with the transfer of *control* of some amount of money; (40c) does not describe any transfer of control over things, but is surely related to *human interest*; (40d) can be seen as related to domains of human interest and force-dynamics, but not to the spacio-temporal domain and the control domain (Newman 1996:2):

- (40) a. Kim gave Lee a nice birthday present.
(where Kim actually handed over her birthday present to Lee)
- b. Kim gave Lee a nice birthday present.
(where Kim arranged for some money to be deposited in Lee’s bank account)
- c. Kim gave Lee emotional support.
- d. Kim gave Lee a kiss.

Mukherjee (2005) investigates the construction based on corpus data and by doing so tries to present a *usage-based theory* of the ditransitive construction. He explicitly claims that “corpus evidence not only tells us important things about actual language use, but also about the cognitive entrenchment (i.e. speakers’ linguistic knowledge) of ditransitive verbs.” (Mukherjee 2005:2). He reveals that the instances with verbs typically used in the construction (e.g., *give*, *tell*) are relatively high and those with atypical verbs (e.g., *deliver*, *drop*) are extremely low, which means that the typicality of the construction clearly corresponds to the frequency

distribution of the verbs used in the construction. This suggests that our knowledge of the construction, or even constructions in general, is based on the actual frequency distribution and hence usage-based.

From those findings it follows that the construction can be seen as a set with gradience at least in terms of its semantics. This characteristics will be discussed later in following sections on the topic of *constructional polysemy*.

4.2.1 Surface patterns

As mentioned in the introductory section of this chapter, the surface patterns of the ditransitive construction is specified based on the findings of previous study by Du Bois (2003). Based on the observation of natural conversation data, Du Bois (2003) argues that there are several preferences in the realization of argument structures in English. Specifically, he finds that the occurrence of a sentence which has more than one lexical arguments (i.e., argument with a full noun, not a pronoun) is fairly rare. This means that a sentence with two or more arguments such as a simple transitive and ditransitive sentences most often has at most only one lexical argument. Examples are as follows, which are retrieved from the corpus of spoken English, named *Santa Barbara Corpus of Spoken American English* (SBCSAE for short, or sometimes simply SBC: Du Bois et al. 2000, 2003; Du Bois & Englebretson 2004, 2005) and hence are accompanied with document IDs of the corpus in parentheses (underlines and italics original; Du Bois 2003:34-35):

- (41) a. DARRYL: you just damn near *broke* the damn needle there? (SBC:DEATH)
 b. KEN: he *named* like half a dozen viruses, (SBC:DEADLY)
 c. PAMELA: ... but I *miss* my grandmother. (SBC:DEATH)

The preference, called the *Preferred Argument Structure*, is considered to be derived from the cognitive and pragmatic need such as informational management (Du Bois 2003:37): in order to avoid presenting too much information to hearer, a speaker tends to minimize the number of informationally-contentful words he/she uses, resulting in the maximal use of words expressing old or given information, namely pronouns.

Statistically, in English the number of clauses with less than two lexical arguments accounts for more than 90% of all in spoken discourse, and this tendency is found cross-linguistically, as shown in Table 4.1.

Table 4.1: The number of clauses with respect to the number of lexical arguments (Du Bois 2003:35, Table 2 [with slight modifications])

Num. of lexical args.	0		1		2		Total	
Language	N	%	N	%	N	%	N	%
Hebrew	261	(50)	252	(48)	9	(2)	522	(100)
Sakapultek	211	(46)	240	(53)	5	(1)	456	(100)
Papago	430	(57)	307	(40)	22	(3)	759	(100)
English	252	(47)	241	(45)	39	(7)	535	(100)
Gooniyandi	2318	(62)	1305	(35)	114	(3)	3737	(100)

In addition to the number of lexical arguments in clause, there is another preference on argument realization, which is about the *positions* of the arguments. As exemplified in (41a-41c), if a clause with two or more arguments has one lexical argument, the argument tends to occupy the *direct object*, as opposed to the subject, position of the clause (Du Bois 2003:35). The same holds true for, importantly, the *ditransitives*, which has three arguments, and therefore a ditransitive clause tends to have pronominal subject and indirect object, as exemplified below (underlines and italics original; Du Bois 2003:37):

- (42) a. JIM: ... he's gonna *send* me those forms, (SBC:BANK)
 b. NATHAN: ... Will you *pass* me some of that tea please. (SBC:ZERO)

In consequence, it is concluded that the typical examples of the ditransitive construction exemplify the sequence [VERB PRONOUN LEXICAL-NP].¹⁾ Moreover, given that the LEXICAL-NP is usually realized as the sequence of an article and a noun in this order, the sequence can be more specified, resulting in [PRONOUN VERB PRONOUN ARTICLE NOUN], which is a leading candidate for a surface pattern of the ditransitive construction. In the next subsection a corpus-based study of the ditransitive construction based on the sequence is presented.

However the categories VERB and NOUN are open-class and hence should not be included in a surface pattern. As a result, the following sequence is considered to function as a surface pattern of the ditransitive construction:

¹⁾It is often the case that the subject is not realized or expressed as seen in, for example, *to*-infinitive clauses (e.g., *I went there to give him a present.*) or in gerunds (e.g., *They will start asking me a question.*), so the subject position is excluded from the surface pattern.

(43) [... PRONOUN ARTICLE ...]

The behavior of this pattern itself is, rather irregularly, not analyzed here, but a related research will be presented in the next chapter (5.2.4).

4.2.2 A corpus-based research (Yoshikawa 2010)

Yoshikawa (2010) investigates exemplar-based nature of the English ditransitive construction based on corpus research. He uses a simple sequence of part of speech [VERB PRONOUN ARTICLE NOUN], which is, as discussed above, considered to be represent a surface pattern of the construction, and searches *Corpus of Contemporary American English* (COCA, Davies 2008-) using the sequence as a query.²⁾ Here it should be pointed out that in the sense that it uses part of speech information, the study cannot be seen as truly exemplar-based: it clearly violates *anti-abstractionist thesis*. This point will be discussed later.

The purpose of Yoshikawa's (2010) research is to represent polysemous character of the construction which is pointed out by Goldberg (1995) from an exemplar-based point of view; he pursues a possibility that the polysemy of the construction is derived from a kind of *cluster effect* on exemplars. This means that he assumes that the construction only *looks* polysemous, but the truth is just that the meaning of the construction comes from a set of exemplars and the set used as a source of the meaning differs according to the actual words used in each ditransitive sentence.

For example Goldberg (1995:75) remarks that the sentences in (44) are indeed ditransitive, but have different senses (see (33)):

- (44) a. Joe gave Sally the ball.
b. Joe baked Bob a cake.

The former is considered to have a sense represented as X CAUSES Y TO RECEIVE Z and the latter, as X INTENDS TO CAUSE Y TO RECEIVE Z. Yoshikawa (2010) argues that this difference can be attributed to the *pair* of the verb and the noun in object₂ (in this case *ball* and *cake*) of the sentence because the other parts of the sentence are considered to be far less contributory to the meaning of the whole sentence, and therefore from the collected data he extracted all the pairs of the verb and the noun, represented as $\langle V, N \rangle$ hereafter.

²⁾Note that the query actually used is not the exact sequence. Due to the limitation of using the web-interface of COCA, sequences which are composed only of highly frequent items (i.e., word, lexeme and POS tag) are rejected by the server and therefore the sequence [VERB PRONOUN ARTICLE NOUN] could not be used as a query. As a solution the query was divided into several queries, all of which specified the pronoun such as [VERB me ARTICLE NOUN], [VERB you ARTICLE NOUN] and [VERB him ARTICLE NOUN].

Extraction of significant $\langle V, N \rangle$ pairs

After some frequency filtering to exclude minor *V*s, *N*s and $\langle V, N \rangle$ pairs, for all the remaining pairs Yoshikawa (2010) computes *MI-score*, which is considered to be an indication of the associative strength, and screens out the pairs whose MI-score go above a significant standard, namely three (Cf. Barnbrook 1996:98-100). The complete list of $\langle V, N \rangle$ pairs evaluated as significant is shown in Table 4.2. It can be said that the $\langle V, N \rangle$ pairs generally represent several ditransitive senses in a bottom-up fashion.

Here two things should be mentioned as to the result: one is that there are some pairs which cannot be regarded as representing the ditransitive construction; the other is that the polysemous character the result suggests can be seen as different from that discussed by Goldberg (1995). As for the first point, the pair $\langle study, moment \rangle$, whose ID is 63, can hardly be seen as parts of a ditransitive sentence. Examples are given below (underline added):

- (45) a. He studied me a moment. (COCA:AntiochRev)
 b. Frank studied him a moment, then glanced at Lisa. (COCA:Bk:ColdFire)

The same holds true for the pair 72, that is, $\langle scare, lot \rangle$. Examples are:

- (46) a. It scared him a lot. (COCA:SouthwestRev)
 b. And this world scares me a lot. (COCA:RollingStone)

These are indeed objects of error analysis in EBCG.

Remember that in this research the query is the sequence of part of speech, not that of concrete items. Yet the part [PRONOUN ARTICLE] can be seen as lexically-specified, in that all the possible combinations of the two categories can be specified. Therefore, as mentioned in 4.2.1, the surface patterns crucial here are such sequences as [... me a ...] and [... him a ...], which do not contain the verbs *study* and *scare*. From this it follows that patterns including the verb *study* and *scare* evoke some other construction effects than the ditransitive one.

In fact, with the queries [*study* PRONOUN ARTICLE] and [*scare* PRONOUN ARTICLE] you can only obtain non-ditransitive examples from COCA. Examples are provided below (underline added):

³⁾A nonce verb *gim* may be a product from a parsing error of the phrase “gimme”; perhaps the phrase is parsed as [gim me] and the first part, “gim,” is treated as a verb.

Table 4.2: The significant $\langle V, N \rangle$ pairs in [VERB PRONOUN ARTICLE NOUN] on COCA

id	verbs	nouns	MI	id	verbs	nouns	MI
1	blow	kiss	10.09	39	hand	money	5.00
2	gim ³⁾	break	9.45	40	give	look	4.96
3	hand	copy	9.44	41	lend	money	4.93
4	hand	glass	9.35	42	send	note	4.89
5	hand	note	8.72	43	give	copy	4.80
6	pour	drink	8.55	44	excuse	minute	4.79
7	owe	favor	8.55	45	afford	opportunity	4.71
8	hand	cup	8.40	46	write	note	4.62
9	hand	drink	8.38	47	pour	shot	4.59
10	pour	glass	8.11	48	offer	job	4.58
11	hand	card	7.95	49	ask	favor	4.55
12	fax	copy	7.95	50	offer	chance	4.47
13	hand	key	7.82	51	tell	story	4.40
14	give	kiss	7.25	52	take	while	4.40
15	pour	cup	7.20	53	offer	opportunity	4.18
16	buy	drink	7.19	54	deny	right	4.13
17	hand	piece	7.14	55	show	picture	4.09
18	gim	minute	7.11	56	leave	note	3.97
19	give	chance	7.00	57	deny	opportunity	3.95
20	hand	letter	7.00	58	give	minute	3.91
21	send	copy	6.97	59	leave	message	3.78
22	tell	truth	6.80	60	give	key	3.75
23	write	letter	6.44	61	give	check	3.67
24	give	opportunity	6.33	62	give	couple	3.66
25	send	letter	6.13	63	study	moment	3.61
26	give	break	5.90	64	allow	opportunity	3.56
27	write	check	5.88	65	excuse	moment	3.46
28	give	ride	5.83	66	send	picture	3.45
29	ask	question	5.82	67	give	right	3.41
30	offer	drink	5.75	68	teach	lot	3.31
31	shoot	look	5.70	69	remind	lot	3.29
32	give	shot	5.56	70	draw	picture	3.20
33	send	card	5.40	71	give	lot	3.17
34	send	check	5.30	72	scare	lot	3.16
35	gim	hand	5.30	73	earn	place	3.12
36	send	message	5.29	74	ask	couple	3.01
37	offer	ride	5.26	75	offer	cup	3.00
38	do	favor	5.19				

- (47) a. Dr. Fox studies him a beat. (COCA: Mov: BloodWork)
 b. And Paglia seems to be studying her every move. (COCA: CBS_Sixty)
 c. This is what scares me the most. (COCA: Commentary)
 d. She's scaring me a tiny bit. (COCA: NBC_Today)

It is found that all the sentences gained with queries include noun phrases behaving like adverbs such as “a beat,” “every move” and “the most” after the sequence [*study/scare* PRONOUN]. This may mean combinations with each of the verbs and such nouns form another kind of surface pattern which evokes different construction(s) from the ditransitive one.

Now let us move onto the second point, the difference in polysemous characters. The senses of the ditransitive construction Goldberg (1995:75) specifies were shown in (33), but for convenience they are provided again:

- (48) a. ‘X CAUSES Y TO RECEIVE Z’ (e.g., Joe gave Sally the ball.)
 b. Conditions of satisfaction imply ‘X CAUSES Y TO RECEIVE Z’ (e.g., Joe promised Bob a car.)
 c. ‘X ENABLES Y TO RECEIVE Z’ (e.g., Joe permitted Chris an apple.)
 d. ‘X CAUSES Y NOT TO RECEIVE Z’ (e.g., Joe refused Bob a cookie.)
 e. ‘X INTENDS TO CAUSE Y TO RECEIVE Z’ (e.g., Joe baked Bob a cake.)
 f. ‘X ACTS TO CAUSE Y TO RECEIVE Z at some future point in time’ (e.g., Joe bequeathed Bob a fortune.)

Her way of classification can be characterized as that based on the *relation* between the elements of the concept, represented as the capital letters X, Y and Z, denoted by the construction. For example the first sense, “X CAUSES Y TO RECEIVE Z,” is characterized by the CAUSES relation between X and Y and the RECEIVE relation between Y and Z (graphically displayed in Figure 4.1).

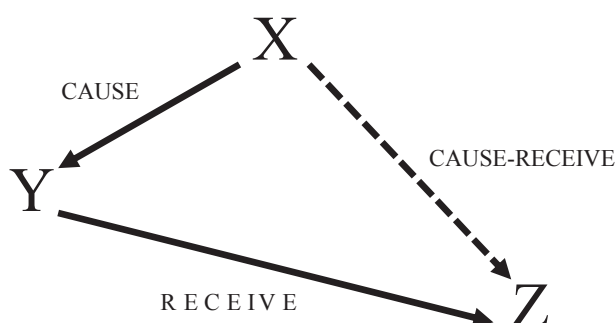


Figure 4.1: Relation network among elements

Relation-based classification clearly does not concern what kind of CAUSE and element is involved. Probably this attitude comes from an assumption of construction as a kind of template with abstract meaning, differentiated from the meaning of a *verb* (see the beginning

part of the section 2.3 and Figure 2.6). However, it is safe to say that there is no guarantee that the polysemous character of the construction can be characterized in a relation-based way. This argument leads to more fundamental problem than the description of construction polysemy: it may be related to the issues on the characterization of construction *per se*.

Simply put, the relation-based characterization can be thought to come from the *verb-construction dichotomy* (see 2.4.2), but this dichotomic assumption seems dubious in some sense, as pointed out in 2.4.3. If so, this line of arguments on the description of constructional polysemy poses a fundamental doubt to the way of theorizing in Construction Grammar. The exemplar-based version of Construction Grammar does not pursue the relation-based nature of the polysemy; more precisely, as argued in the beginning part of this section (4.2.2), the polysemy of the construction may be an *epiphenomena* caused by the characterization of the construction itself in which it is defined as an abstract template without any concrete semantics and forms ([Subject Verb Object₁ Object₂ / X CAUSES Y TO RECEIVE Z]); if it is characterized in a more bottom-up, exemplar-based fashion, such kind of polysemy never appears, but only are there a number of meaning clusters as a result of exemplar clustering.

The type of meaning classification done by Yoshikawa (2010) can be called a *content-based* classification. The two types of classification, relation-based and content-based, make a sharp contrast. The latter can be represented in the way in which the Figure 4.2 displays. Clearly the type of element corresponding to the symbol Z is specified and focused, and the relation is reduced to a simple one (described as an arrow), namely *give*. The content-based characterization, as the name shows, focuses on the content involving in an event or scene, in this case a thing, be it concrete or abstract, which is transferred from one entity, most often a human, to another. This successfully characterizes the difference in meaning according to the direct object noun, in Yoshikawa's (2010) study represented as *N*, as seen below:

- (49) a. He gave me a small question.
 b. He asked me a small question.
 c. He gave me a small gift.
 d. ?He asked me a small gift.

Likewise the contentful character, not the abstract type, of the relation between *X* and *Y* is also involved. Clearly in the following contrast the abstract types of relations among subject referents (*X*), indirect object referents (*Y*) and direct object referents (*Z*) are the same, but the contents of the relation differ:

- (50) a. She gave me a drink.
 b. She poured me a drink.
 c. She bought me a drink.

For all the three examples each Y literally receives Z and hence they should be categorized as members of the type with the most typical sense (48a). However, focusing on what is done by the subject referent X , they are clearly different. It is of course unknown whether the difference should be treated as the matter of constructional meaning or not, and perhaps in the prevailing theories of constructions it would be thought of as the matter of *lexical meaning* of verbs, hence not constructional.⁴⁾



Figure 4.2: A content-based characterization of the ditransitive construction

Formal Concept Analysis of the $\langle V, N \rangle$ pairs

Yoshikawa (2010) also provides a diagram which shows a kind of hierarchical network among the $\langle V, N \rangle$ pairs using a algorithm called *Formal Concept Analysis* (FCA, Ganter et al. 1997, 2005). The diagram is shown in Figure 4.3. It is a kind of graph in a mathematical sense, called a *lattice*. Graph is composed of *nodes* and *edges*, in the way in which nodes are connected by edges. In FCA graph called a *concept lattice* (Ganter et al. 2005:3), each node represents a *concept* in a technical sense. A concept in FCA is defined as a combination of *objects* and *attributes* (Ganter et al. 2005:2). In practical sense, however, there is no qualitative difference between objects and attributes; the FCA algorithm only operates on a *combination matrix* in general, namely a matrix with N columns and M rows ($N \times M$).

In the case of the concept lattice provided by Yoshikawa (2010), the matrix as an input of FCA is that with verbs and nouns which appear in the significant $\langle V, N \rangle$ pair lists presented in Table 4.2. Each box, or a node, of the graph can be seen to represent a *sense* of the construction. in the upper side of the box verbs are listed and in the lower nouns are listed.

⁴⁾Yet a version of Construction Grammar called *Lexical-Constructional Approach* or *Lexical-Constructional-Model* (e.g., ?Iwata 2008) will probably deal with the difference as a matter of constructional meaning and provide proper descriptions of them.

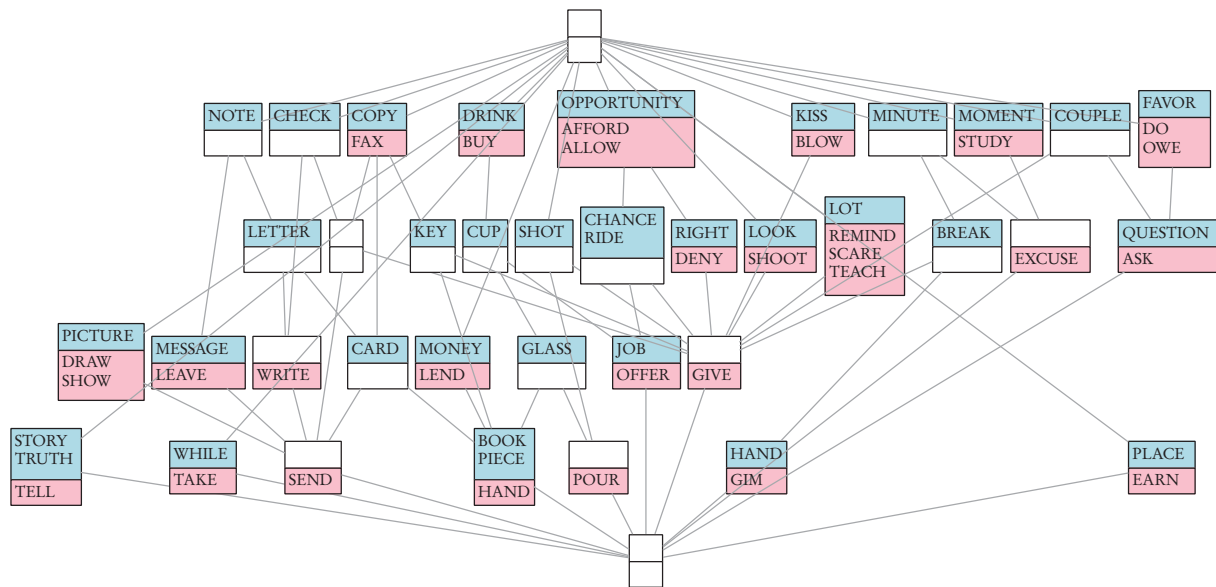


Figure 4.3: A hierarchical network of $\langle V, N \rangle$ pairs in the ditransitive construction

Each hierarchical link, or an edge, represents *subconcept-superconcept-relation* (Ganter et al. 2005:3), in a sense that the upper node is a subconcept of the lower node. The subconcept-superconcept-relation is gained by a simple *inclusion relation*: a concept A is a superconcept of B if and only if A 's attributes are completely included by B 's. In the graph of $\langle V, N \rangle$ pairs attributes correspond to nouns, so, for example, the noun *opportunity* at the topmost middle node is included in all the nodes lower than it if directly or indirectly connected with it (e.g., a node with $\langle right, deny \rangle$, that with $\langle job, offer \rangle$). Just the opposite holds true for the inclusion of objects.

This graph succeeds in visualizing candidates of meaning clusters and their hierarchy. For example, the concept at the top right corner represents the pairs $\langle favor, do \rangle$ and $\langle favor, owe \rangle$ form a cluster, $\langle favor, \{do, owe\} \rangle$; the concept at the bottom left corner represents a cluster $\langle \{story, truth\}, tell \rangle$; there are a large number of concepts which can be seen as subconcepts of the concept $\langle \emptyset, give \rangle$ and $\langle \emptyset, send \rangle$; there are a few concepts which are, as it were, isolated: $\langle place, earn \rangle$, $\langle while, take \rangle$, $\langle \{story, truth\}, tell \rangle$. All of them can be regarded as representing some sort of subsense of the construction in a content-based way.

Problems in Yoshikawa (2010)

As seen so far, the corpus-based research by Yoshikawa (2010) reveals some pieces of truth, but there are a number of problems in it. The most crucial problem is, as mentioned before, that the query used in the research is the sequence of part of speech, which violates the *anti-*

abstractionist thesis. The sequence used in Yoshikawa (2010), however, includes the putative surface pattern presented in (43), namely [... PRONOUN ARTICLE ...], so Yoshikawa's (2010) research can be regarded as revealing partial reality of the construction. The next step is, therefore, to examine how the results change when the sequence changed from [VERB PRONOUN ARTICLE NOUN] to [... PRONOUN ARTICLE ...]. If the results do not change drastically, it is confirmed that Yoshikawa's (2010) findings actually provides some reasonable descriptions of the construction. This task will be done in the next chapter (5.2.4), though in a little different form.

4.2.3 Related study

Goldberg (1995) claims, based on several pieces of evidence, that we need to postulate the existence of the ditransitive construction. One of them is the result of an experiment by her which suggests that a sentence including a novel verb such as (51) can be understood by English native speakers and the interpretation of it is largely stable; she asked non-linguist English speakers what *topamased* meant in the sentence shown in (51) and more than half of them answered that it meant "gave."

(51) She topamased him something. (Goldberg 1995:35)

Further, Goldberg (1995:35-36) remarks that this cannot be explained in terms of the frequency of verbs because according to a word frequency list *give* is not so frequent compared to other verbs such as *tell*, *take*, *get* and *make*. Of course this is the list of overall frequency, not the frequency *in the construction*, and if the frequency count is done within the construction *give* will quite possibly become the topmost verb, but in the above experiment the subjects were not informed that this was an instance of the ditransitive construction; the frequency information in the construction cannot be utilized unless the subjects actively found the sentence to be an example of the construction.

This result suggests that the subjects in some way recognized the sentence as a ditransitive sentence, whether the frequency information of verbs was employed or not. However, there clearly is a problem on this: she does not show any mechanism which enable the subjects to find it ditransitive. In this respect It is quite reasonable to assume that they used a discontinuous lexical sequence, or a surface pattern, [She ... him something], which can be seen as a highly effective evoker of the construction.

In fact in *Corpus of Contemporary American English* (COCA, Davies 2008-) the query *she * him something* (the asterisk * represents a wild card) almost exclusively hits the ditransitive examples. We can get twenty three examples with the query. Some of them are provided below (underline added; document IDs are shown in parentheses):

- (52) a. She gave him something that looked like an empanada. (COCA:ScholasticAction)
 b. She handed him something wrapped in a greasy paper napkin.
 (COCA:FantasySciFi)
 c. ..., and she gets him something from her purse.(COCA:SouthernRev)
 d. She said yes, because she owed him something for his kindness,
 ... (COCA:ParisRev)
 e. One day she told him something was coming between him and the piano.
 (COCA:AmerScholar)

Only one among twenty three hits is not a ditransitive one:

- (53) ..., she thought him something of a fool-too kind-hearted, ... (COCA: VirginiaQRev)

This sentence, however, looks somewhat anomalous and hence it can be seen as an exception.

Table 4.3: The frequency list of verb (forms) in the context [she ... him something]

rank	verb form	freq.	rank	verb	freq.
1	told	3	1	<i>give</i>	5
1	owed	3	2	<i>owe</i>	4
1	gave	3	2	<i>tell</i>	4
4	gives	2	4	<i>hand</i>	3
4	offers	2	5	<i>offer</i>	2
4	hands	2	6	<i>think</i>	1
7	thought	1	6	<i>teach</i>	1
7	tell	1	6	<i>get</i>	1
7	taught	1	6	<i>bring</i>	1
7	owes	1	6	<i>ask</i>	1
7	handed	1			
7	gets	1			
7	brings	1			
7	asked	1			

As for the statistics, the frequency list of what appears in the wild card slot is shown in Table 4.3. As seen in the table *give* is the most frequent. As to verb forms there are two other topmost verbs than *give*, “told” and “owed,” but the former, “told,” is consistently used in somewhat different structure, which is indeed a ditransitive sentence: as shown in (52e), in examples including “told” the word *something* itself is not an object₂, but a *subject* of a complement clause. This may have prevented the subjects of Goldberg’s experiment from associating exemplars of the construction.

4.2.4 Discussions

Here the following questions may arise related to the exemplar-based characterization of the construction:

- (54) a. What about the cases where a non-typical verb is used in the construction?
- b. What about the cases where a non-typical noun is used as the head of the direct object of the construction?
- c. What about the cases where a lexical noun phrase is used as the indirect object of the construction?

Below those likely questions are discussed.

Ditransitive with non-typical verbs

As pointed out in previous studies of the construction, we can find a number of examples of the construction with verbs which are typically not used in the construction such as *kick* and *slide*⁵⁾:

- (55) a. Joe kicked me the ball. (Cf. Goldberg 1995:54)
- b. She slid him the present. (Cf. Goldberg 1995:12)

Moreover, as already seen, the ditransitive construction can be used with some novel verbs such as *skype* or even with non-existent, nonce verbs such as *topamase*, as exemplified below:

- (56) a. She skyped me a photo. [= (16)]
- b. She topamased him something. [= (51)] (Goldberg 1995:35)

⁵⁾On COCA the sequence [*kick* PRONOUN ARTICLE] matches 26 examples and [*slide* PRONOUN ARTICLE] matches 20 examples.

The very character of the ditransitive construction poses a large problem to traditional verb-centered theories of grammar, which assume the structure of a sentence is determined by properties of a head verb (Cf. ?9-23).

Under an exemplar-based framework, the existence of those usages of verbs themselves are not problematic. However, if the source of the constructional meaning of the ditransitive construction is, at least partially, attributed to the combinations of a certain verb and noun in the sequence [VERB PRONOUN ARTICLE NOUN], the status of *verb* in the should be taken into account. In other words, the explanation based on $\langle V, N \rangle$ pairs presented above have to provide some additional accounts for such sentences as in (55) and (56).

EBCG's account is as follows: occurrences of non-typical verbs in exemplars of the ditransitive construction in particular, and perhaps of constructions in general, are analyzable and accountable based on the remaining, probably typical parts such as "me the ball" and "me a photo." Sentences in 55 can, for example, be analyzed based on patterns [... PRONOUN ARTICLE ball] and [... PRONOUN ARTICLE present]. On COCA the former sequence matches 200 examples and of them 63 are used with the verb *give* appearing just before PRONOUN, which is the topmost word at that position. For the latter sequence 266 tokens are found and also in this case the topmost word preceding PRONOUN is *give*, whose frequency is 79. The second-most is *get* and the third-most is *throw*, which occur 61 times and 37 times, respectively. This means that based on the sequences the structures and meanings of those sentence can be successfully described.

Ditransitive with non-typical head nouns of direct object

The same holds true for the head noun of direct object. This may probably be not a problem in the majority of grammar theories, but since in the present theory the role of head nouns of the direct object is emphasized, we have to offer some accounts for examples with non-typical head nouns of direct object. However, just like in the cases with non-typical verbs, non-typical head nouns of direct object can be analyzed based on the remaining, probably typical parts.

Given the arguments so far, it can be generally remarked that EBCG assumes most examples contain at least one sequence which is typical among exemplars of a certain construction, and due to the sequence almost all exemplars can be recognized as an instance of some construction even when there are some atypical parts in them.

Ditransitive with lexical indirect objects

However, there still remains a problem, which is much harder than the previous two. That is about *lexical indirect objects*. Indirect object of the ditransitive construction can, though rarely, be a lexical noun phrase such as *his mother* and *the man*:

- (57) a. Joe handed his mother a letter. (Goldberg 1995:12)
 b. He kicked the man the ball.

Noun is of course an open-class category and hence cannot be utilized in surface patterns under the current exemplar-based framework.

One possible solution here is to make use of the frequent occurrence of *determiners* just before nouns. Determiners are limited in number and considered to form a closed class. Therefore such sequences as [... DETERMINER ... ARTICLE ball] and [... DETERMINER ... ARTICLE present] can be assumed to be utilized. However, there are only a few hits of the sequences on COCA which can be seen to exemplify the ditransitive construction.⁶⁾

Another possibility is that certain familiar nouns can play similar roles to pronouns because of their high frequency. For example, using the sequence [*give* DETERMINER ... ARTICLE] to collect examples of the ditransitive construction including lexical indirect object, words appearing between DETERMINER and ARTICLE position on COCA are found to be such familiar nouns as *child*, *president*, *kid*, *man* and *student* (see Table 4.4). Now the sequences including those frequent nouns such as [... DETERMINER *child* ARTICLE ...] become candidates for surface patterns of the ditransitive construction. Actually [... DETERMINER *mother* ARTICLE ...], which we find in (57a), matches 700 examples on COCA and of them 118 are used with *give*, and [... DETERMINER *man* ARTICLE ...] as in (57b) matches 646 tokens, of which 165 are accompanied with *give*.

There are also cases in which the indirect object noun phrases does not include any determiner, such as follows:

- (58) a. He gave people a chance.
 b. They offer students an opportunity.

In those cases we cannot utilize occurrences of determiner as a part of a surface pattern. However, even in those cases there would be a clue to the recognition. Just like the above

⁶⁾Usually the class determiner includes demonstratives such as *this* and *that*, but since they can also be used pronominally (e.g., *I like this.*), they are excluded when searching the sequence. Specifically the DETERMINER in the sequence actually used in the corpus search only includes articles and *possessive determiners*.

Table 4.4: *X* in [*give* DETERMINER *X* ARTICLE ...]

<i>X</i>	freq.	<i>X</i>	freq.
child	398	son	89
president	240	body	82
kid	201	world	82
man	165	company	80
student	159	people	79
mother	118	daughter	77
guy	112	father	76
boy	91	city	76
room	90	reader	75
team	89	government	72
		⋮	
		total	1584

case, to collect ditransitive examples with determiner-less indirect object, the sequence [*give* ...ARTICLE] where ... is specified as any single word is searched on COCA, which results in finding there are a number of frequently-occurring noun at the ... *position* such as *people* (898 times), *student* (368), *kid* (152), *woman* (148) and *child* (120). Sequences with those nouns also do the same job as the above-examined ones: [... *people* ARTICLE] matches 898 examples accompanying *give* out of 4193 total hits, and [... *student* ARTICLE] matches 1930 tokens, of which 368 are used with *give*.

Given those results suggest that in the case of pronoun's absence there would be some other cues alternatively navigating to the recognition of the construction. However, unfortunately, this is not probably the case. Sentences presented in (55) originally contain lexical indirect objects, as exemplified below:

- (59) a. Joe kicked Bill the ball. (Goldberg 1995:54)
 b. She slid Susan the present. (Goldberg 1995:12)

Indirect objects of both sentences are *proper* nouns. here are actually a large varieties of proper nouns and the word class should be regarded as open. Therefore it is probably impossible to assume the existence of the word class in an exemplar-based way.

There is, unfortunately, any solution to this problem. Probably for those cases some kind of *semantic* approach will be in need. It is safe to say that the distinction between common

noun and proper noun can be recognizable without any knowledge of grammar necessary to learn in school. The distinction is based on highly primitive classification between *type* and *token*; nouns are designed to represent types of things, not tokens, and hence a word class which represents tokens, namely proper noun, should stand out, resulting in their recognizability. If so, we could assume such a sequence as [... NAME ARTICLE ...] of which NAME represents word for tokens, i.e., proper noun.

4.3 The resultative construction

The second construction is the *resultative* construction, which is represented as [Subject Verb Object Xcomp] (Goldberg 1995:3), where Xcomp is a kind of predicate also called the *resultative predicate* (RP) which is either an adjective, a prepositional phrase or a noun phrase. Examples are the following:

- (60) a. I painted the wall red.
 b. She broke it into pieces.
 c. He hammered the metal flat.
 d. Sam talked himself hoarse. (Goldberg 1995:194)
 e. The joggers ran the pavement thin. (Carrier & Randall 1992:217)

As for its semantics, it is described as [X CAUSES Y TO BECOME Z] where X, Y and Z correspond to the subject, the object and the Xcomp, respectively. The part Z is seen as a kind of a *complement* in the grammatical sense and hence regarded as a predicate. In consequence the position can only be filled with an item which functions as a predicate, which means that most adjectives can be placed there, but non-adjectivized participles, locational or directional PPs and most noun phrases cannot (Cf. Boas 2003:to be added).

A large number of studies have been devoted to the construction such as Boas (2003), Carrier & Randall (1992), Jackendoff (1990), Levin & Rappaport Hovav (1991), and Simpson (1983), to name but a few. It can be pointed out that, however, almost all the studies focus on the boundary conditions or the licensing conditions of the construction, namely the conditions which determine whether an instance of it is acceptable or not, so the recognition problem on it has not been investigated in detail (note that Nakatani (2007) argues about the construction in terms of the sentence-processing strategy, which marks a large contrast with the other theory-oriented studies).

From this it follows that there still remains much room for analyzing the construction from the exemplar-theoretic perspective, focusing on the recognition problem of it.

4.3.1 Surface patterns

Boas (2003) presents a list of predicates which can function as the Xcomp of the English resultative construction compiled based on a kind of meta-analysis of the sentences examined by the previous studies of the construction (Boas 2003:15-16). The list contains 51 predicates including both adjectives and prepositional phrases. Here is the list:

- (61) *apart, awake, black, calm, clean, crooked, dead, deaf, dirty, dry, empty, famous, fat, flat, full, hoarse, ill, insane, into pieces, into shape, into the ground, mad, off, open, over the brink, over the edge, over the top, red, safe, shut, sick, silly, sleepy, smooth, sober, soft, solid, sore, stupid, tender, thin, tired, to death, to fame, to insanity, to madness, to pieces, to sleep, to suicide, unconscious, wet*

Though this cannot be completely exhaustive and may contain somewhat problematic items (e.g., the preposition/particle *off*), the exact number of predicates which can work as the Xcomp of the construction should not be far larger than 51, perhaps at most one hundred or smaller.

This suggests that the construction can be captured by the sequence including the specific word or words listed in (61) in proper contexts, though each of the predicates alone cannot play the role. Therefore, in this section, this possibility is pursued mostly based on the data assembled by Boas (2003) and his arguments.

Boas's (2003) findings

Boas (2003:Chapter 5) argues that, based on the data from the *British National Corpus* (BNC) licensing of the resultative construction can be largely attributed to lexical semantics of verbs and for each resultative predicate analyzed by him there is a group of verbs which are strongly associated with the predicate, resulting in forming conventional units composed of a verb and an RP (see Boas 2003:Appendix A). For example, the adjective *dead* is most frequently used with the verb *shoot* as when participating the resultative, as exemplified in (62a), while the PP *to death*, which seems very similar to *dead* as a resultative predicate, most often co-occurs with *shoot*, as seen in (62b). The verbs collocating with the RPs are listed in Table 4.5.

- (62) a. Kim shot Pat dead.

b. Kim stabbed Pat to death.

Table 4.5: Verbs appearing in the resultative construction s with RP *dead* and *to death* in BNC (Boas 2003:130-131)

<i>dead</i>		<i>to death</i>	
verb	freq.	verb	freq .
shoot	408	stab	114
cut	11	beat	74
kill	9	put	44
strike	8	batter	39
stop	6	frighten	34
⋮		⋮	
total	434	total	592

Other verb-RP pairs which Boas (2003) finds to be strongly associated include *<tear, apart>*, *<shake, awake>*, *<wipe, clean>*, *<get, dirty>*, *<suck, dry>*, *<make, famous>*, *<press, flat>*, *<shout, horse>*, *<make, ill>*, *<drive, insane>*, *<cut, into pieces>*, *<get/knock, into shape>*, *<drive, mad>*, *<beat/bite/blow/rip/brush, off>*, *<push/throw, open>*, *<put, over the top>*, *<make, safe>*, *<slam, shut>*, *<make, sick>*, *<tear, to pieces>*, *<put, to sleep>* and *<knock, unconscious>*.⁷⁾ It can be said that the sequence of [VERB ... RP] where the VERB and RP are any of the pair specified in the list can serve as surface patterns of the resultative construction.

Passives and inversions

Boas (2003) does not only provide the statistics of the verb-RP collocations, in which what verbs co-occur with each RP and how frequently they are is presented, but also offers the complete list of the sentence examples of the resultative construction assembled from BNC. The data in the portable document format (.pdf) is downloadable at the publisher's site of the book (<http://web.stanford.edu/group/cslipublications/cslipublications/site/1575864088.shtml>). With this data we can conduct our own quantitative and qualitative analyses of a large number of attested sentences. Here a few findings related to the surface patterns of the construction obtained from the data are provided.

⁷⁾This list is made via somewhat arbitrarily selecting from the tables provided in the Boas (2003:321-340 (Appendix A)).

The list includes 5218 sentences exemplifying the construction, and of them as many as 952 examples are found to be passive sentences, which accounts for almost 20%. Examples are as follows (the sentences are sorted by each RP and indexed with integers within each RP group, so here the examples are presented with ID composed of the RP and the index in parentheses, represented as *RP-index*; underline added):

- (63) a. British athletics may be torn apart by ITV contract. (BOAS:apart-327)
 b. However, the surface could be accidentally wiped clean of its image, forever, as easily as tape. (BOAS:clean-13)
 c. AN OFF-DUTY soldier was shot dead last night in the heart of Loyalist West Belfast. (BOAS:dead-53)
 d. He was standing with his back to her, hands pressed flat on the kitchen table , and he was dragging deep breaths into his lungs. (BOAS:flat-24)
 e. The 70 ton carcass was later cut into pieces and buried. (BOAS:into _ pieces-37)
 f. Almost instantly this time the door was flung open. (BOAS:open-34)
 g. Their mother had been battered to death with a hammer. (to _ death-24)

Interestingly, the passive examples are generally those with strong verb-RP associations such as *tear-apart* (63a), *wipe-clean* (63b), *shoot-dead* (63c) and so on. The verb-RP pairs frequently appearing in the passive version of the construction are listed in Table 4.6.

Table 4.6: The verb-RP pairs frequently used in passives

verb-RP	freq.
<i>shoot-dead</i>	328
<i>tear-apart</i>	64
<i>stab-to _ death</i>	63
<i>throw-open</i>	49
<i>beat-to _ death</i>	38
<i>put-to _ death</i>	30
<i>auction-off</i>	22
<i>crush-to _ death</i>	15
<i>batter-to _ death</i>	13
<i>beat-off</i>	13
<i>burn-to _ death</i>	12
<i>push-open</i>	12

Also interestingly, the same tendency is found in what can be called the *inversion* version of the resultative construction, namely the construction in *Verb RP Object* order, not in *Verb Object RP* order. There are 406 examples (out of 5216) of the construction with the inverted word order, which, if the passives are excluded, accounts for about 10% (406/(5216-952)). Examples are as follows (underline added) and the list of verb-RP pairs frequently used in the inverted word order is presented in Table 4.7:

- (64) a. Israeli soldiers shot dead three Palestinians and injured more than 60 others.
(BOAS:dead-34)
- b. Mr Gorbachev, all the same, managed to beat off this challenge to his leadership.
(BOAS:off-60)
- c. I pushed open the door and bounded in. (BOAS:open-79)
- d. A hot-tempered individual, he early on fought a duel and beat to death an Indian accused of murder. (BOAS:to_death-54)

Table 4.7: The verb-RP pairs frequently used in inverted order

verb-RP	freq.
<i>beat-off</i>	85
<i>push-open</i>	79
<i>shoot-dead</i>	46
<i>throw-open</i>	43
<i>bite-off</i>	18
<i>tear-open</i>	15
<i>rip-open</i>	14
<i>prise-open</i>	12
<i>wrench-open</i>	10

The table shows that the RP *open* is relatively frequently used in the inversions. In fact, of 396 examples of *open* used as the RP of the resultative in Boas's (2003) data, as many as 216 sentences are in the inverted word order, which accounts for more than half of the whole usages.

From these facts it can be pointed out that some, not small part of the resultatives include the [VERB RP ...] sequence, as opposed to [VERB ... RP]. Furthermore, the fact that the number of the passive examples is large suggests that the sequence [*be* ... RP] may also function as a surface pattern of the resultative construction.

Noun-RP sequence also matters

Closer look at the data assembled by Boas (2003) from BNC, it is found that there are some strong associations between a certain noun and an RP. The combination of a noun and an RP has never been focused in the major studies on the resultative construction probably because of the explicit or implicit *verb-centered* or *predicate-centered* view of grammar, in which the central component of the structure of a sentence/clause is assumed to be the predicate of it. Under the exemplar-based framework, however, such an assumption cannot be adopted unless we can find any exemplar-based characterization of the predicate-centered features of sentence structures. On the contrary, under our framework, if there is any piece of exemplar-based evidence which shows items other than the predicate in a sentence does some job as to the sentence structure, those pieces of evidence are actively utilized.

There are a few sequences of *noun-RP* which are found to occur frequently in the resultative sentences. They are *door-open*, *door-shut*, *head-off* and so on, as exemplified below (underline added):

- (65) a. Mark jerked the door open and turned to face the American eyes hardened with anger. (BOAS:open-37)
- b. I snatched his sliding door open and went back into the arctic cold where I punched the button to summon the lift. (BOAS:open-229)
- c. So saying he went back inside and banged his door shut behind him. (BOAS:shut-5)
- d. Curval clicked the case shut and turned, looking back at DeVore. (BOAS:shut-28)
- e. Then , taking a deep breath , began to bawl its head off. (BOAS:off-52)
- f. We all chat our heads off don't we? (BOAS:off-509)

The statistics of *noun-RP* sequences are shown in Table 6.3. As for the *door-open* sequence , of 396 sentences with the RP *open* there are 41 examples whose direct object is *door* in the non-inverted word order (i.e., the exact sequence of [*door open*] is included), accounting for more than 10%. As for *door-shut*, surprisingly, 74 out of 207 examples of the construction with *shut* include *door* as their direct object noun in this order. These data strongly suggest that those [NOUN RP] sequences can also function as surface patterns of the resultative construction. This possibility is pursued in the next subsection (4.3.2).

Table 4.8: The noun-RP sequences frequently found in the resultative construction

noun-RP	freq.
<i>head-off</i>	104
<i>door-shut</i>	74
<i>door-open</i>	41
<i>eye-shut</i>	31
<i>hand-dirty</i>	24
<i>leg-off</i>	20
<i>ball-off</i>	16
<i>hand-off</i>	13
<i>dust-off</i>	12
<i>pant-off</i>	12
<i>dirt-off</i>	11
<i>leg-apart</i>	11
<i>man-to_death</i>	11
<i>sock-off</i>	11
<i>clothes-off</i>	10

The mosaic-like character of the resultative construction

The above-presented arguments on the surface patterns of the resultative construction suggest that the construction does not form one single type of sentences whose member share the same defining properties, but only is a loose cluster or simply a *mosaic* of partially semantically-related phrases or semi-fixed idioms. This is in fact the conclusion provided by Boas (2003): he claims that the construction is strongly based on the lexical semantics of verbs used in the construction and “the lexical-semantic information associated with a word is to a very large degree conventionalized and can therefore not be predicted on general grounds” (Boas 2003:121).

If this is true, we do not have to capture any generalized properties shared in all the examples of the construction, but only to describe each phrases or semi-fixed idioms, which Boas (2003) calls the *mini-constructions*, independently. Therefore the task for our exemplar-based approach to the construction is to provide somewhat detailed descriptions of each of the major *mini-constructions* which consist of the mosaic of the resultative construction.

4.3.2 Corpus-based research

In this subsection the behaviors of several surface patterns of the resultative construction are investigated in order to capture the nature of the construction in an exemplar-based fashion. The patterns examined are: those including the prepositional phrase RPs such as [... to death] and [... into pieces], those composed of pronouns and RPs such as [... PRONOUN mad] and [... PRONOUN clean] and the noun-RP sequences [... door open] and [... door open]. For all of the patterns to be investigated, verbs are *unspecified*, because, as mentioned above, most studies on the resultative construction (and perhaps other constructions also) put primary importance on verbs. This study, in contrast, tries to present the importance of other parts than the verb of a sentence/clause, which will probably lead to the complementary description of the construction.

The data used here is the *Corpus of Contemporary American English* (Davies 2008-), a large-scaled balanced corpus of, as the name shows, contemporary American English including about 450 million words. The reason for selecting this corpus is that it is the largest corpus of English which is available online for free.⁸⁾

RPs of prepositional phrases

As for the adjectives which are listed in the RP list (61) such as *dead*, *mad* and *open*, there are several possibilities of realization: the prenominal modifier (e.g., *the dead man*), the post-copula predicate (e.g., *He is dead.*), the *Xcomp* such as RP (e.g., *She shot him dead.*), and so on. In contrast, the prepositional phrases participating the resultative construction (e.g., *into pieces*, *to death* and *to sleep*) are considered to be more restricted in terms of its syntactic behavior, though they can also function as several syntactic components other than *Xcomp* of the resultative, such as the postnominal modifier (e.g., *the time to death*) and a part of some larger phrases (e.g., *from birth to death*). Therefore sequences including the prepositional phrases serving as RP, *to death* and *into pieces*, are examined.

As for the former sequence, there are 9865 hits of [... to death] on COCA. Since the number of hits is very large, 100 examples out of them are randomly sampled using the builtin sampling function of COCA's web interface and analyzed. Based on the sampling survey, it is found that 85 out of 100 examples instantiate the resultative construction. The 16 exceptions

⁸⁾Of course the tendency found in this corpus may differ from that in Boas's (2003) data obtained from the corpus of British English because of the difference in variation, but so far no divergent characters between British and American data has been found.

are categorized into two, which are 1) those forming some larger unit such as *close to ...* (66a, 66b) and 2) those in which *death* is a part of some complex noun such as *death sentence* (66c, 66d).

- (66) a. Is that about as close to death as anyone can come? (COCA:CBS_48Hours)
 b. ... about the role played by concentration and forced labour camps (as opposed to death camps) (COCA:HistoryToday)
 c. Where you'll find the harshest drug laws, from extensive jail time to death sentences. (COCA:Atlanta)
 d. Particularly chilling are the revelations about efficient Germans making sure the trains carrying the victims to death camps ran on time. (COCA:Atlanta)

The latter sequence, [... into pieces], matches 648 examples. Of them 205 tokens are accompanied with the verb *cut* within three words before *into* and 170 are with *break*, which means that over the half of the total hits contain either of the two verbs and probably the verbs function as a main verb of each sentence/clause. If we can label as members of the resultative construction the phrases with verbs of, say, *changing*, such as *translate X into Y* and *turn X into Y*, almost all the 648 examples matching the sequence are considered to be the resultative. Rare exceptions are those in which *into* describes a path and goal of motion and *pieces* functions as a kind of *numerical classifier*, as exemplified below:

- (67) a. He had sorted books and rolled glasses into pieces of newspaper, ... (COCA:Ploughshares)
 b. The Mirs drop only water yet they, too, lose small parts or bump into pieces of the wreck during dives. (COCA:AssocPress)

They contain somewhat frequent sequences which are thought to evoke constructions other than the resultative, namely [*roll ... into ...*] and [*bump into ...*] and hence can be ruled out.

In this way the RPS of prepositional phrases themselves are considered to be strongly associated with the resultative construction and therefore function as evoker of the construction.

[... PRONOUN RP]

As conducted in the case of the ditransitive construction presented in the previous section, the sequences including pronouns as the direct object of the resultative clause are investigated.

It is expected that the occurrence of a pronoun just before an RP can successfully eliminate almost all the instances of other constructions than the resultative. The RPs examined here are adjectives whose frequency is high compared to others and not participating other case studies presented below, which are: *mad* and *clean*.

First, let us look at how the sequence [... PRONOUN *mad*] behaves. The sequence matches 923 examples on COCA, but they include a number of cases which cannot be seen as exemplifying the construction. Eliminating other possibilities, the number of remaining examples considered to instantiate the resultative is 589, which is almost equal to the number of its occurrences in Boas's (2003) data from BNC, 147, given the corpus scale difference (COCA is about 4 times larger than BNC).⁹⁾ Examples are the following (document IDs of COCA are presented in parentheses):

- (68) a. Doesn't it make you mad a little bit? (COCA:CNN_Talckback)
- b. The touch was impersonal, and driving him mad all the same. (COCA:Bk:Firelight)
- c. ... but it gets him mad as hell. (COCA:FeministStud)

The list of words appearing just before the sequence is presented in Table 4.9. What most often appear at that position is the verb *make*, which surely participates in the resultative construction, as seen in (68a). The second-most word is, however, not considered to be a part of a resultative clause. In fact it consists of an interrogative sentence whose subject is the pronoun before *mad*, as exemplified in (69a) below. Cases where the clauses matching the sequence instantiate the resultative are only when the word appearing at pre-pronoun position is either of the following three verbs: *make*, *drive* and *get*. In most of the other cases the sequence form other constructions than the resultative, such as interrogatives, the *small clause construction* as in (69b) and (69c), or what can be called the *verb-less copulative sentence* as in (69d).

- (69) a. Are you mad about getting older? (COCA:GolfMag)
- b. They think me mad, but this is better than nothing. (COCA:SouthernRev)
- c. That's the first time I ever saw him mad. (COCA:CBS_Morning)
- d. "You mad? You mad or what? ..." (COCA:NewYorker)

For those cases we can find somewhat familiar sequences which seem to compete with those functioning as surface patterns of the resultative construction. For example, the sequence [...

⁹⁾The figure does not include the occurrences of instances with non-pronominal object such as *drive a man mad* while Boas (2003) data does, which makes the comparison not fair. It is found, however, that the objects of as many as 105 out of 147 examples are pronominal, so there would be almost no problem.

think PRONOUN ...] and [... *see* PRONOUN ...] match 276675 examples and 145428 examples on COCA, respectively, which suggests that the sequences evoke the small clause construction or some others much stronger than the resultative, resulting in the whole sentence/clause recognized not as the resultative.

The only case which cannot be determined whether it exemplifies the resultative or some other constructions is where the verb *have* precedes a pronoun, as exemplified below:

- (70) a. You have somebody mad at you? (COCA:Bk:CubaLibre)
 b. They had nobody mad at them, as far as I knew. (CBS_48Hours)

It is easy to say that in the above examples the word *mad* functions as a *verb*, not an adjective, and hence they are different from the resultative construction, but under the exemplar-based framework we cannot rely on any *a priori* assumption on the existence of part-of-speech without exemplar-based characterization of it.

Table 4.9: *X* in [X PRONOUN mad]

rank	X	freq.
1	make	439
2	be	227
3	drive	95
4	get	55
5	think	14
6	n't	9
6	call	9
8	,	8
9	"	6
10	have	5
	⋮	
total		923

Secondly the sequence [... PRONOUN clean] is examined. The sequence matches 1458 examples on COCA and of them at least more than 220 are considered to instantiate the resultative, which are accompanied with verbs such as *wipe*, *make*, *get* and *wash*. Examples are as follows:

- (71) a. She wiped it clean with a paper towel and pressed it against her chest
 ... (COCA:FantasySciFi)
- b. Do people at water companies put chlorine in the water to make it clean?
 (COCA:ChildDigest)
- c. Get it clean, as in no blue streaks on those patches at all. (COCA:FieldStream)
- d. The next morning I went out to the lake, and washed myself clean. (COCA:Ms)

The situation is almost the same as that of [... PRONOUN mad]. A large part of non-resultative examples (at least 267 out of 1458) are of the small clause construction, as exemplified below:

- (72) a. Why don't the state and cities keep it clean? (COCA:Houston)
- b. Calm down, Hajj, let me clean your wound. (COCA:SouthwestRev)

Those examples can also be eliminated by assuming that the sequences such as [... *keep* PRONOUN ...] and [... *let* PRONOUN ...] compete with [... PRONOUN clean] and they are much stronger as evokers of certain constructions other than the resultative (in fact there are a huge number of examples matching those sequences on COCA).

The major difference from the case of [... PRONOUN mad] is that the word *clean* is somewhat frequently used as a verb, though, as mentioned above, *mad* is also used in that way in some rare cases. This tendency is reflected in the fact that the third- and forth-most words appearing just before pronoun in the sequence are punctuation characters, the *comma* and the *period*, respectively, as shown in Table 4.10. It is estimated that in more than half of the total 1458 hits *clean* is used as the main verb and the pronoun before it functions as the subject, as seen below:

- (73) a. And I clean his room every day. (COCA:PBS_NewsHour)
- b. How do you clean soap? (COCA:EEnvironmental)
- c. Would you clean up for dinner? (COCA:IowaRev)

In most of those cases the words appearing at the pre-pronoun position are conjunctions such as *and* as in (73a) or auxiliaries such as *do* and *would* as in (73b, 73c). Therefore they can also be eliminated by assuming the existence of competing sequences such as [... and PRONOUN] and [... AUXILIARY PRONOUN].

Table 4.10: *X* in [X PRONOUN clean]

rank	X	freq.
1	keep	233
2	help	92
3	,	87
4	.	77
5	wipe	62
6	make	53
7	do	46
8	get	37
8	be	37
10	let	34
	⋮	
total		1458

[... NOUN RP]

As shown above, we can find some strong associations between a certain noun in the object NP and an RP in the examples of the resultative construction assembled by Boas (2003) from BNC and therefore the sequences composed of the noun-RP pair in this order could also function as surface patterns of the construction. Here the possibility is examined.

The first sequence is [... door open], of which we find 2413 examples on COCA. A 100-sample survey results in 38 out of 100 examples instantiate the resultative construction. The majority of the non-resultative tokens are either of the small clause construction (74a, 74b) or the *perception verb* construction (e.g., *hear the door open*: 74c). The number of hits of those competing constructions are 40 and 13, respectively.

- (74) a. I will turn off the light and leave the door open. (COCA:Ploughshares)
 b. He held the door open for her. (COCA:Bk:DiscordsApple)
 c. She heard the front door open and close. (COCA:Bk:JaneVowsVengeance)

Although the number of non-resultative examples is much larger than that of the resultative, we can successfully rule out those examples. The small clause construction is considered to be evoked by sequences such as [*leave* ... open], [*hold* ... open] and [*prop* ... open]. For the perception verb construction, unlike the small clause construction it is somewhat hard

to assume such sequences as [*hear* ... *open*] and *see* ... *open* play roles of surface patterns because of their rarity, but lexically less specified sequences like [*hear* ARTICLE ...] can be seen as competitors. Examples neither of the small clause construction nor the perception verb construction include what can be called the *accompanying situation* with construction, such as the following:

- (75) a. Grace was upstairs with her door open again, ... (COCA:Bk:BloodInWater)
 b. What I just cant stop thinking about, Marc Klaas, is a car that is running with a door open and ... (COCA:CNN_Grace)

They can also be eliminated by assuming that the sequence such as [... *with* ARTICLE ... *open*] does the job of a surface pattern (and actually the sequence matches 613 examples on COCA).

The next sequence investigated is [... *door shut*]. In short, the results are much alike that of the [... *door open*] presented just above. The number of total hits is 787 and the majority of them can be seen as instances of the resultative. Just like the sequence with the RP *open*, the non-resultative examples include those of the small clause construction (76a), of the perception verb construction (76b) and the accompanying situation *with* construction (76c). They can be, likewise, ruled out by assuming the existence of competing patterns with *hold*, *hear* or *with*.

- (76) a. "Hold that door shut," I said, ... (COCA:CSMonitor)
 b. Then I heard the door shut. (COCA:AntiochRev)
 c. She was on the phone with the door shut. (COCA:Bk:ConfessDangerous)

There is, however, one notable difference from the results of [... *door open*], which is the existence of examples whose subjects are *door*, with *shut* functioning as the main verb. Examples are as follows:

- (77) a. The door shut behind her. (COCA:BkJuv:Charmed)
 b. I'd been at this a while before a door shut down below. (COCA:Bk:Savage)
 c. He went into the bedroom, and the door shut quietly, ... (COCA:Analog)

For most of those examples it is possible to assume that sequences such as [*and* ARTICLE ...] and [*before* ARTICLE ...], or more generalized one like [*CONJUNCTION* ARTICLE ...], compete with [... *door shut*] and therefore are not recognized as instances of the resultative, so they can probably also be ruled out.

4.3.3 Related study

As for the mosaic-like character of the construction, Goldberg & Jackendoff (2004:535-536) also remark that the construction should be regarded as a “family” of constructions, not a single one, though a generalization over the members of the family is actually possible. They argue that the construction shows a wide range of variety in terms both of syntax and of semantics. Syntactically, the resultative predicate, as mentioned above, can be either an adjective, a prepositional phrase or a noun phrase, and the status of the object is also divergent, ranging from the what is called *selected* object (78a) in the sense that the object is selected by the *selectional restriction* of the verb, an idea based on the traditional verb-centric view of grammar, through the *unselected* object (78b), to the so-called *fake-reflexive* (78c):

(78) a. Bill broke the bathtub into pieces.

(cf. Bill broke the bathtub.)

b. They drank the pub dry.

(cf. *They drank the pub.)

c. We yelled ourselves hoarse.

(cf. *We yelled ourselves. / *We yelled Harry hoarse)

(Goldberg & Jackendoff 2004:536)

What is remarkable in their arguments is that, while trying to make syntactic, semantic and aspectual generalizations, they admit that the productivity of the construction cannot be captured based on one single condition or characterization (Goldberg & Jackendoff 2004:558). They actually analyze a number of idiosyncratic cases and show that the productivity of construction differs almost on a case-by-case basis, with special reference to Boas’s (2003) arguments. Although their arguments are not exemplar-theoretic, but are, say, *construction-based*, we can say that they and we, together with Boas, think along the same line, at least in terms of the analyses of actual data, namely, in a descriptive, if not theoretical, aspect.

4.3.4 Discussions

As seen above, the resultative construction can be described as a mosaic of semi-fixed idioms or mini constructions, as discussed by Boas (2003) or Goldberg & Jackendoff (2004). However, there is at least one problem which has yet to be discussed so far but is important for the exemplar-based characterization of the construction. The problem is on the RPs not included in Boas’s (2003) list, which is examined below.

Given the conventional nature of the resultative construction, it is crucial for us to find what words or phrases can function as RP. This is in fact a methodological problem, but, at least in principle, we have already reached a conclusion in terms of, say, the *policy* of the exemplar-based characterization of the construction. In short, we assume that all the RPs can be analyzed in the same way that those from the 51 list by Boas (2003) are analyzed above: for each of them there would be a verb which is strongly associated with it as seen in the case of *shoot* and *dead*, or some *NOUN-RP* sequence which can serve as a surface pattern of the construction just like the sequence [... door open]. If so, the methodological problem is also solved, because EBCG already provides a methodology to describe constructions based on lexical sequences.

For example, the adjective *senseless* can also function as an RP, as exemplified below (obtained from COCA):

- (79) a. Ever so brotherly, but Cixi wanted to beat him senseless with his cane.

(COCA:Bk:PrinceStorms)

- b. It had knocked him senseless, but thankfully no more. (COCA:Bk:Black)

On COCA, the sequence [... PRONOUN senseless] matches 76 examples and almost half of them (33) are used with the verb *beat*, which gives us an insight that the sequence [*beat* ... senseless] can function as a surface pattern of the resultative construction. In fact, the sequence successfully leads us to the 73 instances of the construction on COCA, with the gap between *beat* and *senseless* set from zero to two (hence including those examples with *beat* adjacent to *senseless*).

There are also cases where a certain prepositional phrase not included in Boas's (2003) list works as an RP. In that case it is expected that, just like the cases of *into pieces* and *to death*, the prepositional phrase itself would evoke the resultative construction. For example, as will be seen in the next section, especially in 4.4.3, there are a number of prepositional phrases which can be used as RP headed by the preposition *into*, such as *into action*, *into play* and *into practice*. The sequence [... into action] matches 1603 tokens in COCA and randomly sampled 100 examples of them show that the phrase can actually function as an evoker of the construction, as exemplified below:

- (80) a. Love had moved someone into action. (COCA:People)

- b. For some reason, the hopelessness of the gesture galvanized Lily into action.

(COCA:BkSF:AlchemistsJournal)

4.4 The caused-motion construction

The third construction is the *caused-motion* construction. The construction is represented as [Subject Verb Object Oblique] where Verb is a non-static verb and Oblique is a directional phrase (Goldberg 1995:152). Examples are the following:

- (81) a. They laughed the poor guy out of the room.
 b. Frank sneezed the tissue off the table.
 c. Mary urged Bill into the house.
 d. Sue let the water out of the bathtub.
 e. Sam helped him into the car.
 f. They sprayed the paint onto the wall. (Goldberg 1995:152)

As for its semantics, it is described as X CAUSES Y TO MOVE Z, in which X, Y and Z correspond to the subject, the object and the noun phrase in the oblique prepositional phrase, respectively.

The construction is strongly associated with some types of verb which in many cases lexically denote the caused-motion scene, such as *put* and *throw*, but not limited to that. In fact many examples presented above in (81) include unusual verbs as caused-motion ones (e.g., *laugh* and *sneeze*).

4.4.1 Surface patterns

In many cases an instance of the construction contains a directional preposition such as *to*, *into*, *through* and so on, and hence this type of preposition can be seen as an indicator of the construction. However, the existence of a directional preposition itself cannot be either a necessary or sufficient condition of instantiating the caused-motion construction. For example, the former two of the following sentences actually contain directional prepositions (*through* and *to*) but do not instantiate the construction, and, contrastively, the latter two do not contain any directional prepositions but can be said as examples of the construction:

- (82) a. I read it through the night.
 b. We enjoyed it to some extent.
 c. He threw the paper in a trash can.
 d. Sam pushed him within arm's length of the grenade. (Goldberg 1995:158)

In addition, as mentioned above, the existence of one of the leading candidates for determining the sentential meaning, namely the *verb* of a sentence, cannot be a necessary nor sufficient condition of the caused-motion construction.

This fact is mentioned by Goldberg (1995) as a piece of evidence that the construction exists. The fact itself, however, cannot be seen as evidence of the existence of the construction. As repeatedly discussed above, also in this case she completely ignores possibilities of construction evoking with multiple-word sequences. In this section, therefore, focusing on the directional prepositions, combinations of them and some verbs or nouns are investigated as to whether they can work as surface patterns of the caused-motion construction.

Ambiguities related to the construction

First we should take into account the issues of *structural ambiguities* related to the caused-motion construction. Let us compare the following sentences:

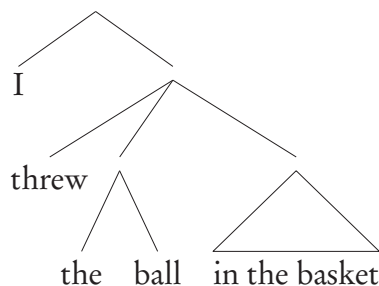
- (83) a. I threw the ball in the basket.
 b. I threw the ball in the park.
 c. I threw the ball in the room.

The first and second sentences, forming a minimal pair, are usually interpreted as having different structures, which can be represented by bracketing as follows:

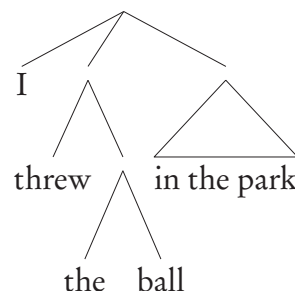
- (84) a. I [threw [the ball] [in the basket]].
 b. I [[threw [the ball]] [in the park]].

Otherwise it can also be represented using tree diagrams as follows:

- (85) a. I threw the ball in the basket.



- b. I threw the ball in the park.



This means that the prepositional phrases “in the basket” and “in the park” differ in their structural functions: the former is, technically, verb phrase-internal but the latter externally

modifies the verb phrase. Semantically, the former describes the throwing of the ball into the basket, while the latter denotes a scene in which the throwing of the ball occurred in the park; put otherwise, “in the basket” represents a *goal* or a *direction* of the throwing but “in the park” represents a *location* in which an action occurs.

The third sentence (83c) is more problematic. It is concerned with both of the two structures described above, namely, structurally ambiguous: the prepositional phrase “in the room” can be interpreted either as a verb phrase-internal element describing the target area of the throwing action, or as a verb phrase-external modifier denoting the location where the action occurred.

Moreover, there is another kind of ambiguity. Let us compare the following sentences, the former of which is equal to (83a):

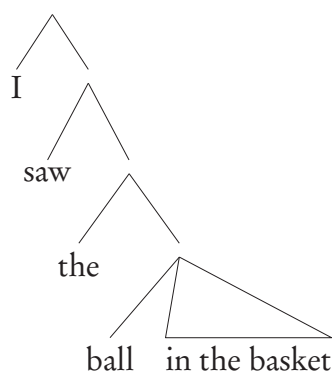
- (86) a. I threw the ball in the basket. (= 83a)
 b. I saw the ball in the basket.

In this case the prepositional phrase “in the basket” is shared by the two and hence they do not differ in that regard. Yet they form a minimal pair in terms of the verb, *throw* and *see*. In fact the latter sentence is usually analyzed radically differently from the former, represented as follows by bracketing:

- (87) I [saw [the ball [in the basket]]].

and by a tree diagram:

- (88) I saw the ball in the basket.



The prepositional phrase in the latter sentence is used *adjectivally*, as opposed to *adverbially*, that is, modifying the noun *ball*.

In view of the two types of structural ambiguity, the caused-motion construction cannot be said to be evoked unless we see the combination of a verb and a prepositional phrase such as *(throw, in the basket)* and *(see, in the basket)*. It is true that structural ambiguity nevertheless remains in such a sentence as (83c), but the ambiguity is also problematic for a language user in the sense that we should solve the ambiguity when interpreting it, and hence it is enough for EBCG to explain that kind of sentences with two or more competing surface patterns.

However, in some cases the combination of a verb and a prepositional phrase is also useless as an evoker of the construction. For a sentence with somewhat unusual verb for the construction such as (81b), the combination does not make any sense because the combination is quite rare or in the worst case has never been experienced. For convenience the sentence (81b) is presented again:

(89) Frank sneezed the tissue off the table. (=81b) (Goldberg 1995:152)

Inversely, there are also cases where the prepositional phrase plays a role of an evoker of the construction, such as the following:

(90) I threw the ball into the park.

This sentence may sound somewhat bizarre, but, if acceptable, it should be interpreted as an instance of the caused-motion construction.¹⁰⁾ The sentence quite resembles the sentence presented in (83b) but the latter is almost unambiguously interpreted a non-caused-motional. The difference is the preposition (*in* vs. *into*) and hence it may be assumed that the prepositional phrase *into the park* always evokes the caused-motion construction regardless of the verb of a sentence or any other factors.

This is, however, not the case:

(91) The window offers a glimpse into the park.

In this case the prepositional phrase “into the park” functions as an adjectival phrase modifying the noun preceding it, *glimpse*. This may be caused by the sequence [*glimpse into ...*], which is considered to evoke a noun phrase construction.

¹⁰⁾The sentence becomes less bizarre if the object noun phrase is changed to *a stone*:

(i) I threw a stone into the park.

Physical and metaphorical motions

Both of the two types of ambiguity described just above are structural, but, there is yet another kind of ambiguity, namely *semantic* ambiguity. When some expression or pattern is semantically ambiguous, the expression or pattern has two or more different interpretations, though it is not structurally ambiguous.

The caused-motion construction is also concerned with this type of ambiguity. Let us compare the following sentences, which share the sequence [He put the ball into ...]:

- (92) a. He put the ball into the basket.
b. He put the ball into play.

The former sentence describes a scene in which the ball was moved into the basket, that is, a physical or literal caused-motion scene. The latter, however, describes a somewhat different scene: the state, not the location, of the ball was changed to an active state, which can be labeled as a *metaphorical motion*.

The important thing here is that the latter type of sentence denoting a metaphorical motion cannot be seen as an instance of the caused-motion construction, but one of the *resultative* construction. In fact Goldberg (1995:81-89) argues that the resultative construction can be seen as a *metaphorical extension* from the caused-motion construction, under the name of the *Unique Path Constraint* (Goldberg 1991), which states that, regarding the resultative predicate as a metaphorical *location* and the metaphorical change of location as a metaphorical *path*, there cannot be two or more paths within a single clause.¹¹⁾ At the same time, however, the two constructions are not two different realizations of one and the same single construction, but are two different constructions existing independently but connected by the *Metaphorical Extension Link* (Goldberg 1995:87), as briefly introduced in 2.3.1.

This semantic type of ambiguity is certainly problematic for EBCG because they seem quite difficult to be separated based on surface patterns. However, there is a possibility that if we specify the prepositional phrases which evoke the resultative construction when placed in that context, then we can identify the remainder as examples of the caused-motion construction. Given that the resultative construction can be seen as a set of relatively fixed expressions

¹¹⁾The exact characterization by her is the following:

If an argument *X* refers to a physical object, then no more than one distinct path can be predicated of *X* within a single clause. The notion of a single path entails two things: (1) *X* cannot be predicated to move to two distinct locations at any given time *t*, and (2) the motion must trace a path within a single landscape. (Goldberg 1995:82)

without large variability resultative predicates, specifying the phrases does not seem impossible. This possibility is pursued later in 4.4.3.

Relation between the object and the oblique nouns

Generally, the relation between the object and the oblique noun also has effects on determining the construction. The object noun in a caused-motion sentence can be either a human or a thing, and the oblique noun can be a human or a place, which means that there are four possible patterns of the relation between the two nouns: human-human, human-place, thing-human and thing-place.

It is predicted that the thing-human pattern prefers construction somewhat different from the caused-motion construction, called the *transfer-caused-motion* construction, such as *He gave it to me*, while the human-human pattern (e.g., *He take me to her*) seems not to have strong preference; when some place is involved, however, whether the object is a human or a thing, the simple caused-motion construction is preferred (e.g., *He brought it to the house* and *He brought me to the house*).

Actually the results of corpus search show the preferences as predicted (see Table 4.11 to 4.14). The corpus used here is COCA and the types of noun are specified as follows:

- (93) a. HUMAN object: {*me, you, him, her, us* }
 b. HUMAN oblique: {*me, you, him, us* }
 c. THING: *it*
 d. PLACE: [ARTICLE {*home, place, house, room, area, space, position, site, station, corner, spot, apartment, flat, location, residence* }]

As for the oblique noun the pronoun *her* is eliminated because of its possessive usage. The nouns included in the PLACE sequence is selected from COCA's synonym list of the noun *place*.

It can be pointed out that most verbs listed in Table 4.11 are considered to be those also involving the ditransitive construction, represented by the topmost word *give*. This clearly suggests that these verbs, perhaps except the fourth one, *explain*, and the ninth, *put*, participate in a somewhat different construction from the caused-motion construction. The construction may be what is called the *transfer-caused-motion* construction (we will see the construction somewhat in detail in 5.2.3 especially in the subsection titled 160).

This tendency can further be confirmed by seeing other prepositions than *to* because *to*

Table 4.11: [X THING *to* HUMAN]

rank	X	freq.
1	give	2076
2	hand	584
3	show	464
4	explain	438
5	send	342
6	bring	305
7	do	257
8	put	210
9	sell	184
10	leave	161
	⋮	
total		7397

Table 4.12: [X HUMAN *to* HUMAN]

rank	X	freq.
1	bring	144
2	and	125
3	pull	124
4	to	103
5	draw	84
6	give	83
7	,	69
8	from	58
9	send	53
10	have	50
10	that	50
	⋮	
total		2181

Table 4.13: [X THING *to* PLACE]

rank	X	freq.
1	take	25
2	make	18
3	bring	9
3	get	9
5	send	8
6	return	5
7	sell	4
7	move	4
9	carry	3
9	apply	3
9	attach	3
9	transfer	3
	⋮	
total		159

Table 4.14: [X HUMAN *to* PLACE]

rank	X	freq.
1	take	199
2	lead	50
3	bring	48
4	drive	32
5	with	22
6	send	14
6	get	14
8	carry	12
9	show	11
10	walk	10
	⋮	
total		590

is the only preposition involving the transfer-caused-motion construction. In fact the search using the sequence [X THING *into* HUMAN] results in providing almost no verbs listed in Table 4.11 (see Table 4.15).

Table 4.15: [X THING *into* HUMAN]

rank	X	freq.
1	drill	14
2	drum	4
3	beat	3
4	get	2
4	pound	2
6	put	1
6	read	1
6	take	1
6	blow	1
	⋮	
total		41

The importance of specific nouns

In addition to the prepositions, specific nouns also play a role as an evoker of the construction when combined with some preposition. For example, the preposition *off* cannot be an evoker of the construction in itself, as seen in the examples below:

- (94) a. A charter boat plies the waters off the coast of Hoonah (COCA:NPR_Morning)
 b. ... I saw a flash off the window ... (COCA:ABC_PrimeTime)
 c. The charging system should have charged the batteries off the engine.
 (COCA:Analog)
 (Examples from COCA; document ids are shown in parentheses; underline added)

None of these exemplify the caused-motion construction: the prepositional phrases headed by the preposition *off* (e.g., “off the coast” in (94a)) encode the *location* in which some events occur, not the *direction* or *path* of any motion.

However, as seen in (81b), the preposition can be used in the caused-motion construction with typically non-causative verbs such as *sneeze*, which means that the caused-motion sense cannot be attributed to neither the verb, the preposition, nor the combination of them. In this regard, Kuromiya (2010) provides an important finding that the noun in the prepositional phrase headed by *off* plays a role. The noun he specified is *table*. He found that all the sentences including sequence [NOUN *off the table*] in *British National Corpus* (BNC) exem-

plify the caused-motion construction. More interestingly, Kuromiya (2010:409) provides the following verb-less sentences including the phrase *off the table* which clearly have the caused-motion sense:

- (95) a. Elbows off the table. (BNC:CH8)
 b. Feet off the table! (BNC:KD5)

Kuromiya (2010:410) explains this fact as follows: the phrase *off the table* is strongly associated with the verb *take* and hence would probably “incorporate” the meaning of *take*, or the whole verb phrase *take ... off the table*. In fact, of 40 examples of the phrase he found in BNC, as many as 16 followed the verb *take* (Kuromiya 2010:413). This reasoning is fairly exemplar-theoretic.

Identifying the surface patterns

In summary, it can be concluded that there is no general rule or pattern for evoking of the caused-motion construction. Within the framework of EBCG, the construction should be investigated almost on a case-by-case basis, moving the focus among the preposition or prepositional phrase, the verb, the object noun phrase, and any pairs of triples of them.¹²⁾

There is, however, a possible way to obtain some generality in terms of specification of the surface patterns, though it is only a *passive* approach in the sense that it does not tell us what specific sequences play roles as evokers of the construction, but, conversely, it specifies the sequences which function as the evokers of any other constructions such as the resultative construction and the transfer-caused-motion construction. In other words, if we can eliminate all the ambiguities related to the evokers of the construction discussed above, it will indirectly lead to the specification of the construction in an exemplar-based fashion.

In many cases the ambiguities can be solved by focusing on some competing sequences strongly associated with other constructions than the caused-motion. For such a sentence as shown in (96b), the sequence [*investigation into ...*] can be seen as a leading candidate for the competitor, which would be an evoker of a noun phrase construction *investigation into X*; For the sentence presented in (96c), the sequence [*...into pieces*] probably competes with such sequences as [*...kick the ball into ...*] but is stronger than them, and hence the whole sentence is recognized as a resultative. Inversely, without those competing sequences, such pattern as [*... kick ... into ...*] as seen in (96a) function as evokers of the caused-motion

¹²⁾Somewhat generalized arguments and quantitative analyses will be presented in the next chapter (5.2.3).

construction.¹³⁾

- (96) a. He kicked the ball into the park.
 b. He conducted an investigation into the park.
 c. He kicked the ball into pieces.

4.4.2 Goldberg's (1995) puzzle

In order to show the effectiveness of EBCG's methodology, this section focuses on one single sentence exemplifying the caused-motion construction, which seems highly problematic for the exemplar-based theory of constructions. The sentence is presented by Goldberg (1995:158) as an example of the caused-motion construction whose source of the "motion" meaning cannot be explained with any approaches other than constructional one:

- (97) Joe squeezed the rubber ball inside the jar.

She argues that the sentence (97) is unambiguously interpreted as an instance of the caused-motion construction, meaning that the rubber ball is caused to move inside the jar in a "squeezing" manner. Its caused-motion meaning, however, cannot be reduced either to the meaning of the verb included in it, *squeeze*, or to the meaning of the preposition *inside*, and hence the meaning should come from elsewhere than the words composing the sentence.

Her solution to this problem is the reduction of the meaning to the construction it instantiates, namely the caused-motion construction: the sentence has that meaning because it is an instance of the construction (Goldberg 1995:158-159). This kind of reasoning is quite common in her argument for the existence of constructions, as seen in 2.3.1 and 2.4.4.

From the perspective of EBCG, however, the meaning of the sentence (97) should be analyzed without attributing it to the constructional meaning of the caused-motion construction. If we can find effective surface patterns in it which unambiguously specify appropriate exemplars with the caused-motion meaning, the above raised puzzle by Goldberg (1995) can be solved in an exemplar theoretic fashion. This is actually possible.

¹³⁾ In this connection the following sentence may remain ambiguous (if acceptable):

(ii) He kicked an investigation into the park.

Preliminaries

Before going to the analysis part, one preprocessing should be done here. The sentence (97) includes a noun phrase containing a complex noun, *rubber ball*, which may make the available data highly sparse if queried with the exact sequence. Therefore the part *rubber ball* is simplified into *ball*, hence the sentence investigated becoming the following:

(98) Joe squeezed the ball inside the jar.

The behavior of *squeeze*

Now let us move on to the analysis. First, the verb *squeeze* is, somewhat contrary to Goldberg's (1995) argument, frequently used in a caused-motion sense. In fact, if we search COCA (Davies 2008-), we can find a lot of examples of the verb in the caused-motion construction:

- (99) a. So you just squeeze it into the pastry bag, and ... (COCA:NBC_Today)
 b. I squeezed myself into my old homecoming dance gown and ...
 (COCA:AmerScholar)
 c. Nothing short of magic could squeeze it through a traffic jam, whatever the cause.
 (COCA:Analog)
 d. ..., he squeezed her to his chest and ... (COCA:Moment)

Statistically, there are 1257 hits for the sequence [*squeeze* PRONOUN], which can be seen as the surface pattern of the transitive construction with the verb *squeeze*, and it is estimated that about at least 12 percent of them are used as the caused-motion construction, given that 164 out of 1257 hits of the sequence are followed by the prepositions *into*, *through* and *to*, which can be regarded as evidence of the caused-motion usage of *squeeze*, as seen in the example above.

The question here is whether examples including *squeeze* can be seen to have the caused-motion meaning if it is combined with the preposition *inside*. On COCA, there are 42 hits for the sequence [*squeeze* (...) *inside* ...]. The interval between *squeeze* and *inside* is between zero and three words, which covers the objective noun phrase in the original sentence (97), namely *the rubber ball*. Though intransitive usage is dominant, the obtained list is filled with sentences with motional meaning; 25 out of 42 are intransitive (100) and 15 are transitive (101):

- (100) a. Maggie squeezed back inside the van. (Bk:WithVengeance)

- b. Then they squeezed inside and the door was sealed again in their wake.

(Analog)

- (101) a. He squeezed himself inside the cell and found a munchkin-size bathroom ...
(COCA:Bk:DayAtonement)
- b. Punch a small hole in one and squeeze the gas inside onto a burning match.
(COCA:KansasQ)
- c. He came in a gipsy taxi and squeezed his bag inside as she opened the door.
(COCA:KansasQ)

There are also two examples of the *way* construction:

- (102) a. Tanner squeezes his way inside. (COCA:Mov:Metropolis)
- b. Several people recognize Hank as they squeeze their way inside.
(COCA:Mov:EightScenes)

This result suggests that the sequence [*squeeze (...) inside ...*] can function as an evoker of some motion construction, causative or intransitive. As for the two examples of the *way* constructions, as will be seen in the following section (4.5), it is assumed that the lexically specific sequence [POSS *way*] (e.g., *his way*) does some job to evoke the *way* construction and hence the caused-motion construction is not evoked.

The pattern [... ARTICLE ball inside ...]

Additionally, the pattern [... ARTICLE ball inside ...] can also be considered to play a role. Here the definite article *the* is generalized into an article. The pattern matches 52 tokens on COCA. Examples are shown below:

- (103) a. Normally you want to get the ball inside and get fouled, ... (COCA:NYTimes)
- b. North Atlanta had the ball inside the Southside 5-yard line late in the game ... (COCA:Atlanta)
- c. I was throwing the ball inside. (COCA:Atlanta)

Not all the examples are instances of the caused-motion construction, many can be seen as denoting caused-motion scene, as seen in (103a) and (103c) above.¹⁴⁾

¹⁴⁾See <http://corpus.byu.edu/coca/?c=coca&u=8387&k=5885> for all the concordance lines

This tendency can be confirmed by seeing the words appearing just before the sequence. That place is the position which in many cases verbs occupy. The list of the words is presented in Table 4.16. The table shows that the verb *get* quite frequently appears at that position, which amounts to about 30% (= 16/52). As seen in (103a), examples with the verb *get* at the position are considered to be caused-motional. In addition, the second most word *pound* can also be seen as functioning in the same way as *get*. Examples are the following:

- (104) a. ... rookie Ron Dayne must pound the ball inside, ... (COCA:Denver)
 b. Texas never pounds the ball inside anyway. (COCA:Atlanta)

The sum of the occurrences of *get* and *pound* amounts to almost to the half of the total hits (= 25/52).

Table 4.16: Words preceding [ARTICLE *ball inside*]

rank	word	freq.	rank	word	freq.
1	get	16	8	throw	1
2	pound	9	8	with	1
3	have	3	8	in	1
4	pump	2	8	inside	1
4	run	2	8	into	1
4	dribble	2	8	keep	1
4	work	2	8	line	1
8	control	1	8	of	1
8	dump	1	8	pass	1
8	feed	1	8	place	1
8	take	1	8	play	1
8	that	1			

On the relation between caused-motion and intransitive motion

Here it may be a problem that, as seen above, the intransitive usage is dominant for the sequence [*squeeze* (...) *inside*], because the construction in question is the *caused-motion*, not the intransitive motion, construction. This is, however, assumed to be non-problematic.

Goldberg (1995:78) argues that the intransitive motion construction is connected with the caused-motion construction by what she calls the *subpart link*, which connects constructions if “one construction is a *proper subpart* of another construction and exists independently”

(Goldberg 1995:78, emphasis original). This means that if her characterization is correct, those two constructions are completely compatible with each other in terms of semantics.

From this it follows that within the framework of EBCG, given the *non-contradictory condition* presented 3.3.9, exemplars of the intransitive motion construction and those of the caused-motion construction should be successfully summated into one, because they are non-contradictory at least in terms of their meaning.

The remaining problem is as to the process of unification. The meaning obtained by summing the exemplars associated with the pattern [*squeeze (...) inside ...*] should be unified with that obtained by summing those with the pattern [... ARTICLE *ball inside*]. Fortunately this problem can also be solved naturally within the framework of EBCG.

As described in 3.3.10, unification is done with four steps: meaning decomposition, proper alignment, wildcard insertion, columnwise unification. For convenience the whole process of unification is presented again (= 24):

- (105) a. **meaning decomposition:** for a pair of $\langle x = \mathfrak{S}(\sigma_i), y = \mathfrak{S}(\sigma_j) \rangle$, decompose each of the two into parts, resulting in an array $x^* = [x_1 x_2 \dots x_n]$ and $y^* = [y_1 y_2 \dots y_m]$, respectively;¹⁵⁾
- a'. the decomposition should be done so that the number of overlapping parts between those from the two, i.e., $|x^* \cap y^*|$, would become maximum;
- b. **proper alignment:** sort the order of the one whose size is smaller so that the order of its elements would maximally correspond to the other;
- c. **wildcard insertion:** for each array, insert a special element \bullet so that every element of the two would completely correspond;
- c'. the special element \bullet functions as a kind of *wildcard*, and therefore the logical product of something with \bullet always returns the thing itself, that is, $\forall x(x \wedge \bullet = x)$;
- d. **columnwise unification:** for each pair of i th element in x^* and y^* , that is, $\langle x_i, y_i \rangle$, make the logical product of the two, $z_i = x_i \wedge y_i$, and return the newly obtained array of the products, $z = [z_1 z_2 \dots z_l]$;
- e. sum up z and return the sum as $\mathfrak{U}(\sigma)$.

The key is the step (105a). The meaning obtained from the pattern [*squeeze (...) inside ...*], represented as $\mathfrak{S}(\sigma_1)$, should be decomposed so that its parts maximally overlap those

¹⁵⁾Recall that $\mathfrak{S}(\sigma_i)$ and $\mathfrak{S}(\sigma_j)$ represent the summated meanings

obtained from [... ARTICLE *ball inside*], represented as $\mathfrak{S}(\sigma_2)$. In this step (105a) the decomposed parts of $\mathfrak{S}(\sigma_1)$ should include a part corresponding to the *caused* argument, namely the object referent of the caused-motion construction, because $\mathfrak{S}(\sigma_2)$ contains the part explicitly denoting the caused argument, that is, [ARTICLE *ball*].

This operated meanings is conceptually represented as follows:

- (106) a. [([someone causes]₀) [something]₁ [(to) move(s) inside]₂ [somewhere]₃
[by squeezing (it)]₄]
b. [([someone causes]₁ [some ball]₂ [to move inside]₃ [somewhere]₄]

The first part of (106a) indexed with 0 is an optional segment. As is clear from the representation, if the former is unified with the latter, the caused-motional meaning should result.

Furthermore, some may pose a question as to the particle usage of *inside*, as opposed to prepositional. In fact many examples with *inside* presented above (e.g., (103a), (103c)). Syntactically, a preposition and a particle are indeed different: the former requires an object but the latter does not. Semantically, however, their difference is not so clear. Particle usage of *inside* in most cases denotes a relation between something or some action and some place, boundary, container or the like, the latter of which can be specified contextually. For example in the case of the sentence presented in (103a), the particle *inside* can be seen as describing the relation between the ball and a field of some ball game.

The pattern [... *inside the jar*]

The remaining part is the sequence [*inside the jar*], which involves the ambiguity between a prepositional phrase construction and some other dependent constructions such as the caused-motion construction and noun phrase constructions. As for the sequence, it should be concluded that the sequence does not evoke any specific construction, judging from examples of the sequence on COCA:

- (107) a. So you put the candy inside the jar and ... (COCA:CBS_Early)
b. Michelle reached inside the jar, ... (COCA:Bk:MissingPieces)
c. Tape the strips inside the jar,... (COCA:Parenting)
d. ..., which were stored in water inside the jar.... (COCA:Antiques)

If fact there are only 14 hits for the sequence on COCA and the sequence is associated with a wide variety of constructions, resulting in the quite weak power of evoking. Therefore, the meaning obtained from this part should, if any, resemble the following:

(108) $\llbracket \text{something or someone is or moves inside the jar} \rrbracket$

As a result, the three meanings obtained from $\llbracket \text{squeeze (...) inside} \rrbracket$, $\llbracket \dots \text{ ARTICLE ball inside} \rrbracket$ and $\llbracket \dots \text{ inside the jar} \rrbracket$ are unified into the meaning like the following:

- (109) a. $\llbracket (\llbracket \text{SO causes} \rrbracket) \llbracket \text{STH} \rrbracket \llbracket (\text{to}) \text{mv(s) inside} \rrbracket \llbracket \text{SWH} \rrbracket \llbracket \text{by squeezing (it)} \rrbracket \rrbracket$
 b. $\llbracket \llbracket \text{SO causes} \rrbracket \llbracket \text{the ball} \rrbracket \llbracket \text{to mv inside} \rrbracket \llbracket \text{SWH} \rrbracket \bullet \rrbracket$
 c. $\llbracket \bullet \llbracket \text{STH/SO} \rrbracket \llbracket \text{is/mvs inside} \rrbracket \llbracket \text{the jar} \rrbracket \bullet \rrbracket$
 $\downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow$
 d. $\llbracket \llbracket \text{SO causes} \rrbracket \llbracket \text{the ball} \rrbracket \llbracket \text{to mv inside} \rrbracket \llbracket \text{the jar} \rrbracket \llbracket \text{by squeezing it} \rrbracket \rrbracket$

This unified meaning is equated with the meaning of the verb phrase of the sentence (98), *squeezed the ball inside the jar*.

4.4.3 Patterns with the preposition *into*

As sporadically shown in 4.4.1, for the caused-motion construction the preposition *into* can be seen as an important lexical indicator which is strongly associated with the construction. This section, therefore, investigates the behavior of the preposition as an important part of the sequences functioning as the evokers of the construction.

Preliminaries: with or without a definite article

First let us examine the list of resultative predicates by Boas (2003) presented in (61). The list contains a few prepositional phrases, especially with the typically direction-coding ones, *to* and *into*:

- (110) *into pieces, into shape, into the ground, to death, to fame, to insanity, to madness, to pieces, to sleep, to suicide*

What is remarkable here is that except the third one, *into the ground*, almost all the phrases do not include any articles, simply composed of the preposition and a bare singular or plural noun. Not included in the list, the prepositional phrase in the above example (92b), *into motion*, also consists of a preposition and a bare noun.

This suggests that the evokers of the resultative construction can be identified as prepositional phrases with bare nouns. Of course there are exceptions such as *into the ground* mentioned above, but if the exceptions can also be specified, there remains no problem with that respect. Leaving the exceptions aside, however, let us investigate how the results change according to the presence of the article if we search the sequences in corpora.

Corpus data

In order to know the general tendency, the sequence [... PRONOUN into the/–the ...] is examined, because it can be seen as a surface pattern of the caused-motion construction, if not evoking the resultative. Searching the sequences on COCA, the following sentences are obtained:

(111) [... PRONOUN into the ...]

- a. He passed me into the arms of Grandma Pat. (COCA:Storyworks)
- b. She slid them into the basin. (COCA:SouthernRev)
- c. I put him into the bed. (COCA:LiteraryRev)
- d. She has turned me into the biggest wimp at home. (COCA:Ebony)
- e. The Angel family has finally made it into the middle class. (COCA:Bk:MigrantSouls)

(112) [... PRONOUN into –the ...]

- a. I've a mind to haul you into court for it. (COCA:Bk:SPQRX)
- b. My wife actually put it into practice by ordering them, (COCA:AmSpect)
- c. Bring it into Los Alamos. (COCA:PBS_Newshour)
- d. We brought him into our apartment. (COCA:AmSpect)
- e. After reading it, he led me into a back room. (COCA:FantasySciFi)

As seen in (111a–111c), many examples containing the former sequence instantiate the caused-motion construction, while some are considered to exemplify the resultative constructions as seen in (111d) and (111e). Those exceptions can be characterized with specific verbs typically encoding change-of-state senses such as *turn*, *make*, *transform* and *convert*, or some fixed expressions different from either the caused-motion construction or the resultative construction such as *make it into*. Put differently, the pattern [*turn* ... into ...] can be seen as an evoker of the resultative construction, and the pattern [*make it into* ...], as an evoker of what can be called the *make it into* construction. On the other hand, it is found that the examples with the latter sequence generally describe metaphorical motion scenes. However, the number of exceptions are relatively large (e.g., (112c–112e)). This point will be discussed later.

Table 4.17 and 4.18 show the words appearing just before the sequence [PRONOUN into the/–the]. They clearly show that the sequence [X PRONOUN into the] prefers verbs encoding physical motion scenes when used in the context (e.g., *throw*, *bring* and *take*), while

the sequence [X PRONOUN into \neg the] has a somewhat different preference. What is remarkable here is the topmost item in Table 4.18, *turn*. It can be said that the verb lexically encodes the change-of-state scene.

Table 4.17: [X PRONOUN into the]

rank	X	freq.
1	throw	589
2	bring	556
3	get	545
4	take	515
5	follow	480
6	make	479
7	lead	416
8	put	413
9	push	265
10	toss	264

Table 4.18: [X PRONOUN into \neg the]

rank	X	freq.
1	turn	3677
2	get	2250
3	put	1818
4	make	1166
5	take	1058
6	transform	991
7	bring	910
8	throw	824
9	talk	806
10	lead	634

The *into VERB-ing* construction

Examples matching the sequence [... PRONOUN into \neg the] can be categorized into several constructions including the resultative. The most remarkable one is what can be called the *into VERB-ing* construction, or the *into-causative* construction, such as follows:

- (113) a. My cousin talked me into coming back out. (COCA:CSMonitor)
 b. Anna deludes herself into believing she has chosen Charles. (COCA:StudShortFic)

This construction can be partially specified with the sequence [... PRONOUN into ...ing] and therefore is not so problematic for the exemplar-theoretic approach.

Caused-motions without a definite article

As mentioned above, there are a number of examples which can be regarded as caused-motional but match the sequence [... PRONOUN into \neg the ...] (e.g., (112c)). Those “exceptions” can be categorized into the following three types:

- (114) a. a proper noun follows *into* (e.g., (112c));
 b. a possessive determiner follows *into* (e.g., (112d));
 c. an indefinite article (*a, an*) follows *into* (e.g., (112e)).

Let us set aside the first condition involving proper nouns for now because the behavior of proper nouns is really problematic, as briefly discussed in 4.2.4. The sequence occurring in the context of the second condition seems to work as an evoker of the caused-motion construction, so the sequence [... PRONOUN into POSSESSIVE ...] could be seen as the evoker. The third one, however, does not lead us to the caused-motion construction, as seen below:

- (115) a. ..., he had worked himself into a rage. (COCA:AmerScholar)
 b. ..., the words had the desired effect of startling them into a brief silence.
 (COCA:BkSF:AlchymistsJournal)

They are examples of the resultative construction.

Although the number of the examples exemplifying the resultative construction accompanied with an indefinite article at the position is so small,¹⁶⁾ it seems that their existence does not allow us to assume the sequence [... PRONOUN into INDEF-ARTICLE ...] works as an evoker of the caused-motion construction. However, also in this case the nouns appearing in the *into*-phrases will probably do some job to make us recognize the whole clause as instances of the resultative. In fact, many examples of the resultative construction if the phrase *into a rage* and *into silence* are searched on COCA:

- (116) a. ..., it tends to whip you into a rage when you run into opposition.
 (COCA:TownCountry)
 b. And that, the plaintiffs claim, then threw him into a rage with respect to Nicole Brown Simpson. (COCA:CNN_Talkback)
 c. But little things would send Barbara into a rage. (COCA:CBS_48Hours)
 d. ... Duncan and Robinson could surround Shaq and frustrate him into a rage, ...
 (COCA:WashPost)
- (117) a. The authoritativeness of Ed's voice jarred Iliana into silence.
 (COCA:Bk:GeographiesHome)
 b. My scream of submission shocked the theater into silence. (COCA:Bk:BlackSilk)
 c. I would have liked to smash it into silence. (COCA:Raritan)

¹⁶⁾A sample survey with one thousand examples of [... PRONOUN into INDEF-ARTICLE ...] on COCA resulted in finding only two examples of the resultative construction if not used with such caused-motion-oriented verbs as *turn* and *make*, all of which are shown just above.

Therefore it can be concluded that the sequence [... PRONOUN into INDEF-ARTICLE ...] also functions as an evoker of the caused-motion construction, because the above examined exceptions, i.e., examples which match the sequence but exemplify the resultative construction, can be eliminated when we take into account other competing sequences including *rage* or *silence* in *into*-phrases.

With or without a definite article, revisited

The final problematic context is one presented in (114a): a context in which a proper noun follows *into*. What is problematic here is that the category of proper noun can hardly be identified in an exemplar-theoretic way, because, unlike pronoun and article, the category is open and there is no formal feature shared by all the members of it.¹⁷⁾ Therefore we cannot utilize the occurrence of proper nouns as a sequence functioning as any kind of surface patterns.

However, fortunately, there seems to be no example containing a proper noun immediately after *into* which exemplifies the resultative construction unless the verb of the whole clause is not a member of change-of-state-oriented ones such as *turn*, *make* and *transform*. In this respect, the existence of a definite article may no longer be a good indicator of the caused-motion construction: with or without a definite article after *into*, if the *into*-phrase or the sequence including the phrase does not evoke any other constructions than the caused-motion, the sequence [... PRONOUN into] will probably evoke the caused-motion construction.

The following contrasts are illustrative of this point:

- (118) a. Some equipment in Building 776 was so radioactive that workers
 hailed it into Room 127, (COCA:Denver)
 b. So we'll either have to haul it into the room, ... (COCA:LiteraryRev)
- (119) a. ..., they've been turning it into the place they've always wanted,
 ... (COCA:WashPost)
 b. ..., you can get us into the place, ... (COCA:Analog)

The former contrast shares the sequence [... *haul* it into ...] and the noun *room* following the sequence but only differs in the existence of a definite article before *room*. Clearly, both examples are instances of the caused-motion construction. In contrast, the latter contrast consists of examples of two different constructions, though sharing a sequence, [... PRONOUN

¹⁷⁾Of course we can almost unambiguously identify proper nouns as such by the first capital letter when written.

into the place]. The example including the verb *turn* (119a) exemplifies a kind of the resultative construction, but the other one (119b) is an instance of the caused-motion construction. The difference in the pronoun appearing just before *into* does not make difference, because both of the former contrast contain *it* at the same position but instantiate the caused-motion construction.

4.4.4 Related studies

Though he argues in terms of the resultative construction, Boas (2003:260-277) provides a similar account for the mechanism of licensing such a novel instance of the caused-motion construction as (120):

(120) Tom sneezed the napkin off the table.

His explanation can be labeled as *analogy-based* in the sense that he attributes the licensing mechanism to the existence of more basic and conventional expressions functioning as the base of analogy from which novel instances are made. For example, as to the sentence (120) the following sentence is considered to be its model (Boas 2003:264):

(121) Tom blew the napkin off the table.

To be precise, his argument concerns about the *verb*, not about the whole sentence or clause, and focuses on the frame-semantics associated with the verb. He therefore calls such a verb as *blow* which functions as the base of analogy a *source verb* (e.g., Boas 2003:268).

As mentioned in 4.3.1 Boas's (2003) argument has many things in common with that of EBCG and the argument above also seems highly similar to the exemplar-theoretic conception of grammatical construction. However, there is at least one crucial difference between his and EBCG's arguments: his argument is not based on concrete exemplars. 260-277 only claims that there is a base of analogy for any novel instances of a construction and the base is a verb which is conventionally used with a syntactic frame seen in the construction in question such as [NP V NP PP], an is associated with a frame semantics fitting the semantics of the construction in question such as [X cases Y to move Z].

In other words he does not seem to have any concerns with the on-line process of grammatical constructions. Even if the syntactic and semantic properties of the sentence or the verb which is the candidate for the analogical base fit the construction in question, there is

no guarantee that we can access the analogy base when we process any novel instances of the construction. Put simply, unless we can access such a sentence as (121) when encountering a novel expression like (120) based on any perceptible information, we cannot utilize the sentence as the base of analogy. In this sense we have to say the analogy-based explanation given by Boas (2003) is crucially different in terms of the *cognitive-realist* thesis of EBCG.

4.4.5 Discussion

As we have seen, the caused-motion construction cannot be easily captured in an exemplar-based way. In this section, the very difficulty of the exemplar-based specification of this construction is discussed and by doing so the true nature of the construction is tried to be investigated.

Ambiguity resolution

As intensively discussed in 82 and 91, ambiguity resolution is highly important for the specification of the caused-motion construction. The ambiguities are three-fold, two of which are structural and the other of which is semantic. In order to specify the caused-motion construction by means of surface patterns, we have to eliminate the possibilities of evoking other related constructions, namely the *locational* PP construction (122b), which has not been mentioned above and hence will be discussed below, some noun phrase construction such as *investigation into X* (122c), and the resultative construction (122d) or more specific phrases such as *make it into* construction (122e). In some cases the transfer-caused-motion construction discussed in 92 (122f) may also be involved.

- (122) a. He kicked the ball into the park. [= (96a)]
 b. He kicked the ball in the park.
 c. He conducted an investigation into the park. [= (96b)]
 d. He kicked the ball into pieces. [= (96c)]
 e. He made it into the park.
 f. He gave the ball to me.

Problem with the locational-PP construction

The arguments provided in this whole section tell us that constructions other than the locational PP construction such as (122b) can be specified by surface patterns which considered to

compete with those evoking the caused-motion construction and hence the ambiguity would not be problematic. Therefore the only concern at present is how to solve the ambiguity between the locational PP construction and the caused-motion construction. We can clearly see the ambiguity in the contrast between sentences (122a) and (122b) above because they form a minimal pair, differing only in the preposition appearing just after *the ball*, so let us examine the behavior of the sequence around the preposition *in* and *into*.

As seen in 4.4.3, the sequence containing *into* can be said to have strong tendency to evoking the caused-motion construction if it also includes a pronoun just before *into*. In contrast, this is not the case with *in*. Statistical analysis and more detailed discussion on this contrast between *into* and *will* be given in 5.2.3, so here the behaviors of larger sequences are investigated, which are [... in ARTICLE park] and [*throw* ... in ...].

As an approximation to the specification of its behavior, first, verbs accompanying the sequence is examined. On COCA (Davies 2008-) there are 3230 hits of the sequence [in ARTICLE park]. The results are obtained by searching verbs with POS tag which occur within three words before the sequence [in ARTICLE park]. Of 3230 examples and of them as many as 428 are used with the verb *be*, but they include auxiliary usages of *be* and hence are set aside in order to avoid the risk of double-count. Other frequently collocating verbs than *be* are *play* (62), *walk* (61), *sit* (56), *see* (48), *find* (45), *take* (41), *live* (39), *meet* (35), *go* (32) and *run* (29) (the figures in parentheses are their frequencies). Examples are the following:

- (123) a. They could still play pickup games in the park and eat at the same lunch table.
(COCA:Highlight)
- b. Gradually, we began walking his dog in the park every evening ...
(COCA:Cosmopolitan)
- c. She would have sat with him in the park, ... (COCA:Bk:SearchingTinaTurner)
- d. They saw a bird in the park. (COCA:USAToday)
- e. Later, they found a bench in the park and watched families walk by.
(COCA:Salmagundi)

Almost all of the 10 topmost verbs presented above are strongly associated with the locational-PP construction as seen in the examples just above. The only verb which is often used in the caused-motion construction is *take*, but examples found in this research do not contain any instances of the construction:

- (124) a. Perhaps they take a walk in the park or drive in a car or whatever.
(COCA:NPR_Sunday)
- b. The first part of every lesson took place in the park, ... (COCA:EnvironEd)
- c. A businessman takes a break in a park in central Tokyo. (COCA:Futurist)

Over the half of the examples (22) with *take* are those of *take a walk* as in (124a) and most of the others exemplify idioms such as *take place* (124b) and *take a break* (124c).

The results suggest that the sequence [...in the park] has strong association with the locational-PP construction, so other things being equal, the sequence will probably evoke the construction, as opposed to the caused-motion construction. This reasoning, however, cannot be validated without examining other seemingly-competing sequences which may evoke the caused-motion construction, and the leading candidate for the competitor is [... *throw* ... in ...].

It seems hard to get any approximation to the behavior of the sequence [... *throw* ... in ...] and therefore for the sequence a sampling-based research is performed, in which a one-hundred sample is extracted from the results obtained in the search of the occurrence of *within* three words after *throw*.¹⁸⁾ The number of the total hits is 6760. Of the 100 randomly-sampled examples, at least 39 are found to be caused-motional, which strongly suggests that the sequence has somewhat high association with the construction. The examples of the caused-motion construction are presented in (125) and those of some other constructions including the locational-PP construction are in (126):

- (125) a. The thief threw the necklace in the water. (COCA:RangerRick)
- b. Laurie snapped off her rubber gloves, threw them in the trash, ...
(COCA:Bk:Blindsight)
- c. I didn't know if you were going to throw coffee in my face or what?
(COCA:PBS_Newshour)
- (126) a. There will be people in the crowd who will throw themselves in front of tanks ...
(COCA:USNWR)
- b. Higuera's fastball topped out at 85 mph, and he threw 68 pitches in five innings ...
(COCA:Atlanta)

¹⁸⁾The sample is selected using the random sampling function provided in the web interface of COCA.

- c. She throws it down in disgust as if her eyes have been burned.
(COCA: Mov: Legally Blonde)

Exploring larger sequences

It should be mentioned that, however, most of the examples above found to exemplify the caused-motion construction contain one of the following three prepositional phrases: *in the river*, *in the trash* and *in one's face*. This may mean what evokes the construction is not the sequence [... *throw* ... *in* ...] but the prepositional phrases or the combinations of them such as [... *throw* ... *in the river*]. In fact, when searching what appears just before the sequence [PRONOUN *in the river*] and [PRONOUN *in the trash*] it is found that the verb *throw* is at the top. As for the former sequence the number of total hits are 83, of which as many as 32 examples include *throw*. The second most is *dump*, which has a highly similar sense as *throw*, but occurs only 6 times. As for the latter sequence 165 examples are found and 65 of them are used with *throw*. The second most verb is *toss*, which occurs 32 times, and then *put* follows (occurring 17 times).

In contrast, searching the sequence [PRONOUN ... *in one's face*] results in finding that the topmost item appearing just before it is *see*, whose frequency is 87 out of 305 total hits. *Throw* is the second most one, which occurs 31 times. It can be said, unfortunately, that the meaning of the phrase *see X in one's face* resembles that of *stare at X* and hence is not similar to that of the caused-motion construction. Of course it is possible to assume that some abstract entity such as the *staring* by someone is caused to move to *X*, but, unlike the usual caused-motion construction such as (125c) what is caused to move is something unstated and not the object of the verb, *X*. In this sense the phrase *see X in one's face* should be concluded to be different from the caused-motion construction.

Ambiguities resolved

However, now fortunately, when the sequence [*throw* ... *in one's face*] is searched with the gap between *throw* and *in*, represented as three dots (...), specified as two maximal words on COCA, it is found that all the examples instantiate the caused-motion construction. The number of the total hits is 142. The same holds true for the sequence [*throw* ... *in the river*], which occurs 54 times, and [*throw* ... *in the trash*], which occurs 129 times.

Now that we can say that the evokers of the caused-motion construction include [*throw* ... *in one's face*], [*throw* ... *in the river*] and [*throw* ... *in the trash*], but such smaller se-

quences as [*throw ... in ...*] would probably be excluded. More generally, for a sentence including the location-oriented, as opposed to the goal-oriented, preposition such as *in*, it is when the somewhat large, lexically highly specific sequences function as the surface pattern that the sentence is judged as instantiating the caused-motion construction; otherwise, if it includes any of the sequences evoking other similar constructions, it is recognized as an instance of one of them, and if not, it is classified as an example of the locational-PP construction.

What exactly is the caused-motion construction?

Based on the conclusion obtained just above, the true nature of the caused-motion can be depicted in an exemplar-theoretic way. Under the framework of EBCG, the construction is characterized, just like the resultative construction, as the group of lexically highly specific constructions surrounding prepositions or propositional phrases. As for goal-oriented prepositions such as *into* and *to*, the construction can be evoked unless the prepositional phrase or the sequence composed of the noun just before the preposition and the preposition itself does not work as an evoker of any other construction such as the *investigation into* construction and the resultative. What happens in the case of location-oriented prepositions is as described just above.

Then future task is to investigate the behaviors of other prepositions which are usually used in the caused-motion construction, namely goal-oriented prepositions or preposition clusters not investigated so far such as *onto* and *out to*, what can be called the *path*-oriented prepositions such as *through*, *across* and *over*, and *source*-oriented prepositions such as *from*, *out of*, and *off of*. Although their behaviors have yet to be examined, the methodology adopted in the research presented in 4.4.3 can also be applied to them and therefore we can say the characterization of the caused-motion construction give just above is testable. For now, given that the through research has not been conducted, testability would be of primary importance.

4.5 The *way* construction

The forth and the final construction investigated in this chapter is the *way* construction. Examples are the following:

- (127) a. Pat fought her way into the room.
 b. Volcanic material blasted its way to the surface.

- c. The hikers clawed their way to the top.

(Goldberg 1995:16)

The structure is represented as [Subject_i Verb *POSS_i way* Oblique] (Cf. Goldberg 1995:199), where *POSS_i* denotes a possessive which has the same reference as that of the subject. The meaning of the construction is characterized as the combination of the *motion* and the *creation of path*. For example the sentence (127a) describes a scene in which Pat created a path to the room by fighting and actually entered into the room (Goldberg 1995:207).

This construction is known to be highly productive, in the sense that a large variety of verbs can be used in this construction, ranging from manner-of-motion verbs such as *inch* (128a), through typically transitive verbs such as *push* (128b), to the intransitive verbs which by no means designate any senses of motion when used in other contexts than in the construction, such as *belch* (128c).

- (128) a. Max inched his way across the ledge. (COCA:?:53)
 b. Max pushed his way through the crowd. (COCA:?:49)
 c. He belched his way out of the restaurant. (COCA:?:202)

4.5.1 Surface patterns

Unlike the constructions investigated so far, the surface pattern of the *way* construction can easily be specified. The pattern is:

- (129) [... *POSS way* PREPOSITION],

where *POSS* means possessive determiner such as *my* and *your* and the *PREPOSITION* is restricted to some subparts of it, which will be discussed later. Here the reason why the sequence can function as the surface pattern of the construction is explained in reference to some previous studies on it including Goldberg (1995). The sequence includes abstract categories such as *POSS*, so the disjunctive version of the pattern is also presented:

- (130) [... *my/your/his/her/its/our/their way to/into/through/across/onto/over/...*]

The importance of *one's way*

Goldberg (1995:199) remarks that the actual motion is implied only when the object is *POSS way*; other semantically similar nouns such as *route* cannot participate in the construction and the possessive determiner cannot be substituted for any article without losing the

motion sense. For example, compared to the sentences in (131), which instantiate the *way* construction, those in (132) do not entail any motion (Goldberg 1995:199):

- (131) a. Frank dug his way out of the prison.
b. Frank found his way to New York.

- (132) a. Frank dug his escape route out of prison.
b. Frank found a way to New York.

The latter pair only denote the situations in which Frank created or discovered a path, whether he actually moved through the path or not.

Prepositions

Goldberg (1995:199) notes that the oblique appearing after *POSS way* codes a directional (but without any verification of the characterization). In other studies on the *way* construction such as Kuno et al. (2004) almost the same thing is pointed out. Kuno et al. (2004:78-79), for example, presents the following contrast and argues that the construction is acceptable when the prepositional phrase expresses a path, as opposed to a location (*italics original*):

- (133) a. John laughed his way *out of the room*. (path)
b. Mike moaned his way *through the tunnel*. (path)

- (134) a. *John laughed his way *in the room*. (location)
b. *Mike moaned his way *in the tunnel*. (location)

Therefore for the *way* construction the preposition appearing just after *way* should be one which can denote a path.

Provisionally the prepositions usable in the construction are thought to be the ones listed below:

- (135) *across, ahead of, along, around, between, down, from, into, onto, out of, over, past, through, to, under, up*

The list is probably not exhaustive, however.

Competing patterns

There are cases where a sentence is not recognized as an instance of the *way* construction even if a sentence matches the pattern presented in (129). One of those competing conditions is the case in which the verb *know* precedes the sequence and the preposition is *around*, namely the sequence [*know* POSS way around]. In fact the following sentence entails neither the creation of a path nor the motion through the path:

(136) He knows his way around town. (Goldberg 1995:200)

This sentence only describes that the person referred to as *He* is really familiar with the town.

The second competitor is the sequence [*be* POSS way PREPOSITION], that is, the sequence in which *be* precedes the surface pattern of the *way* construction. If a sentence matches the competing pattern, it only provides a description of the subject reference being someone's usual practice, as exemplified below:

(137) This is my way of doing that.

The third case in which the *way* construction is not evoked in spite of including the sequence [... POSS way PREPOSITION] is one where the preposition appearing at the PREPOSITION place is *of*, as seen in (137) and below (taken from COCA):

- (138) a. Now do you understand our way of living? (COCA:BkSF:WhenFiveMoons)
 b. But it happened because people changed their way of thinking about Communism.
 (COCA:NYTimes)

The forth and the final competing sequence is [PREPOSITION POSS way PREPOSITION], as exemplified below (taken from COCA):

- (139) a. He's the six-year-old boy who disappeared on his way to school in 1978 ...
 (COCA:CBS_ThisMorning)
 b. Committed fans will go out of their way for a glimpse of a favorite film location.
 (COCA:Newsweek)
 c. Because he is known for his way with a jury, ... (COCA:CBS_EyeToEye)

Except for the third case, all the remaining three can be eliminated based on the condition of pattern inclusion, because they are more lexically-specified than the surface pattern of

the *way* construction. The third case in which the sequence [POSS way of] competes with the *way* construction can also be eliminated by excluding the preposition *of* from the list of prepositions which can function as a part of the surface pattern of the *way* construction.

PP-less variation

Excluding the competing patterns presented above, the sequence [... POSS way PREPOSITION] can work as the surface pattern of the *way* construction. However, there is yet another problem on the exemplar-based characterization of the construction. It is known that there are instances of the construction which do not include any prepositions at the PREPOSITION slot of the sequence. Such instances, which we call the *PP-less variation* of the *way* construction because they do not include prepositional phrases (PP) unlike the other regular members of the construction, can be categorized into two subtypes, the *path-less* type and the *adverbial-path* type. Examples are provided below, of the first type in (140) and of the second type in (141), taken from Kuno et al. (2004:80-81, italics original):

- (140) a. Horticulture was a new course so the staff were *feeling their way* just as much as the new batch of students.
 b. But it quickly hushes up as another group *group its way*, begging for assistance.
- (141) a. But there 's no machine that can tie shoelaces or find its way *home*, ...
 b. Nicole hesitates a moment, then slips to her feet, and slowly wanders her way *northward* for a bit of privacy.

In the examples of the first subtype, the path-less type, there is no specification of the path through which the subject (hence *path-less*). For our exemplar-based explanation the existence of this type of the construction seems somewhat problematic because the good indicator of the construction which is included in the surface pattern of it, the preposition following *way*, is absent.

Fortunately, however, most examples of the *way* construction do include phrases describing paths and the path-less type is quite exceptional. Probably the type is only allowed when the verb is specified either as *feel* or *grope* as seen in (140), so if we add the sequence [*feel/grope* POSS way] to the surface patterns of the construction the type can also be dealt with under the exemplar-based framework. The remaining problem is that we cannot fully know what other verbs than *feel* and *grope* can participate in the path-less type of the *way* construction.

related to the *creation of path* sense of the construction, because the verb *make* lexically has the creation sense.

Under the EBCG's framework the relation can be explained in a straightforward way: the sequence [... POSS way PREPOSITION] is associated with a large number of exemplars containing the verb *make* at the position represented as ... and, because those exemplars have the *creation of path* sense, the meaning obtained via summation of the associated exemplars should have that meaning. Therefore, a sentence instantiating the *way* construction entails the creation of path, regardless of the verb used in the sentence.

In fact, the frequency of the verb *make* is huge. Searching what appears just before the sequence [POSS way PREPOSITION] on COCA, the results presented in Table ?? is obtained. The PREPOSITION slot is specified as either of the following preposition: *across*, *along*, *around*, *back*, *between*, *down*, *from*, *into*, *onto*, *over*, *past*, *through*, *to*, *under*, and *up*. When querying all prepositions are excluded and the result is post-processed so that the verbs *know* and *be* are eliminated. The table shows that *make* occurs 5092 times of the total 17619 hits, accounting for about 29%, more than one-fourth of the total occurrence. This figure is far larger than that provided by Goldberg (1995:206).

Table 4.19: *X* in [*X* POSS way PREPOSITION] on COCA

rank	X	freq.	rank	X	freq.	rank	X	freq.
1	make	5092	11	weave	146	21	navigate	93
2	work	2365	12	claw	141	22	shoulder	69
3	find	2047	13	talk	138	22	grobe	69
4	fight	488	14	wend	114	24	ease	67
5	pick	483	15	elbow	107	24	shove	67
6	push	474	16	inch	103	26	earn	64
7	feel	230	17	eat	102	27	see	57
8	force	221	17	pay	102	28	dance	56
9	wind	199	19	worm	98	29	negotiate	52
10	thread	166	20	buy	97	30	edge	50
							⋮	
							total	17619

4.5.3 Related studies

Kunihiro (1967:121-122) argues that a phrase instantiating the *way* construction which does not contain *make* as its main verb (e.g., *thread POSS way*) still has the meaning of *make POSS way*; in such a sentence the meaning of a verb other than *make* (e.g., *thread*) is *superimposed to make POSS way*. Kunihiro (1967:122) calls the composed meaning via superimposition *stratification*, which is highly similar to the idea of exemplar-theory of constructions. Kunihiro's argument is also similar to that by Boas (2003) presented in 4.4.4 in that the phrase considered to be superimposed to can be identified with the base of analogy.

This means, at the same time, that Kunihiro's (1967) argument cannot be seen as an exemplar-based theory, just like Boas's (2003). Kunihiro (1967) does not base his argument on the concrete exemplars but on abstract phrases such as *make POSS way*. He does not, of course, present any explanation about the on-line process of the construction, either.

4.5.4 Discussion

As seen above, the *way* construction can be specified in an exemplar-theoretic way via surface patterns presented above, and the source of the constructional meaning, the creation of path, can be attributed to the ample examples of the construction including the verb *make* as the main verb. However, there remains a large problem in the characterization presented so far. The problem is: the existence of a semantic variant different from those examined in this section. It is known that the *way* construction has two semantic variants, though structurally the same, namely the *means* interpretation and the *manner* interpretation (e.g., Goldberg 1995:202-203).

The manner interpretation

Goldberg (1995:202) presents the sentence (144) has two different interpretations shown in (145), based on the argument by Jackendoff (1990:211-214):

(144) Sam joked his way into the meeting.

- (145) a. Sam got into the meeting by joking. (means)
 b. Sam went into the meeting (while) joking. (manner)

The latter interpretation (145b) is called the *manner interpretation*, because the verb is interpreted as expressing a manner of motion, as opposed to the means to create a path.

Although Goldberg (1995:202) claims that the manner interpretation is not typical in terms of its low frequency (she found only 40 examples out of 1117 in the corpora she used) and marginal acceptability (she reports that not all English speaker judge the manner interpretation acceptable), if that interpretation is possible at all, we have to be able to explain the source of it in an exemplar-based fashion, in the same way as the other, *means* interpretation. The problem here is that our explanation presented above is based on the fact that the surface pattern [... POSS way PREPOSITION] can specify the exemplars of the *way* construction. The pattern, of course, also matches examples with the manner interpretation and hence we cannot differentiate two interpretation variants based on the surface pattern.

One possible solution to this is to find other competing patterns which can function as an evoker of the manner-interpretation variant of the *way* construction. The putative candidates are the prepositional phrase such as *into the meeting* and sequences including the verbs such as *joke ... into*. Looking at the sentence (144) the former seems to be a leading candidate because the exact phrase *into the meeting* is expected to occur somewhat frequently and hence is possible to work as an evoker. In fact we can find 78 examples on COCA, though the frequency cannot be said to be absolutely large. However, Goldberg (1995:209) also provides the following examples which only allow the manner interpretation, whose prepositional phrases cannot be seen as popular as *into the room*:

- (146) a. [They were] clanging their way up and down the narrow streets ...
 b. ...the commuters clacking their way back in the twilight towards ...
 c. She climbed the stairs to get it, crunched her way across the glass-strewn room ...
 d. He seemed to be whistling his way along.
 e. ...he was scowling his way along the fiction shelves in pursuit of a book.

The verbs used in the sentences, *clang*, *clack*, *crunch*, *whistle* and *scowl*, are also rare, which means that the latter possibility, the sequence including the verb functioning as a surface pattern, is not promising, either.

There is, however, yet another possibility: when the preposition (or adverbial) appearing after *way* is specified as one of a certain type, the exemplars associated with the sequence would drastically change, resulting in the verb *make* falling to lower than the second-most one. If the majority of the associated exemplars do not contain the *creation of path* meaning and, instead, express some manner of motion, the sequence can be seen as an evoker of the manner interpretation variant of the *way* construction. In fact, for the preposition found

in (146a), *up*, corpus data from COCA show that the verb appearing just before *POSS way* is not *make*, but *work*. The sequence [POSS way up] matches 3197 examples, of which as many as 701 contain the verb *work*. *Make* is the second-most, whose frequency is 287. It is still disputable that exemplars with the verb *work* actually has the manner interpretation, as opposed to the means interpretation, but at least the fact that the status of the verb *make* degrades is important.

However, prepositions or adverbials in other sentences than (146a) do not show that tendency. As for *back* as in (146b), the most frequent item appearing before [POSS way back] on COCA is found to be the preposition *on* and the second-most is *make*, whose frequency is 451. The third-most is *find*, occurring 417 times. In this case there is almost no difference between the frequencies of *make* and *find*, so the sequence may behave differently from the general tendency of the *way* construction. In the case of the remaining prepositions and adverbials seen in (146c–146e), however, *make* occurs most frequently at the pre-*POSS way* position. In consequence we cannot attribute the source of the manner interpretation to the sequences including the prepositions and adverbials used in the above sentences.

There is no clear solution to this problem and therefore how to explain where the manner interpretation of the *way* construction comes from is a future task of the exemplar-based theory of construction. For now, though speculatively, a possible hypothesis on that is provided. Perhaps, given the marginal acceptability of the manner interpretation of the construction, that variant is considered to be based on a construction effect different from that of the means interpretation variant, and the construction effect is relatively weak, in the sense that the size of the exemplar set associated with surface patterns of the variant is not so large. Moreover, the source will probably be inconsistent: there would be number of sources which evoke the manner interpretation. In one case the prepositional phrase may do the job and in another some [POSS way PREPOSITION] with the PREPOSITION position specified as some sort of preposition or adverbial may play a role. Related argument is provided below in reference to the phenomenon of *construction* competition.

Notes on construction competition

As manifested in 3.6.3, EBCG does not directly deal with the matters on acceptability. However, in terms of the phenomenon which can be called the *construction competition*, the acceptability problem can be argued under the framework of EBCG, though highly indirectly. The phenomenon, moreover, has some aspects related to the manner interpretation of the

way construction, and therefore it is beneficial for us to discuss the problem in terms of the construction competition.

First, let us look at the following contrast:

- (147) a. *Joe walked his way to the store. (Kuno et al. 2004:73)
 b. *Bill walked his way down the hallway. (ibid)
 c. The novice skier walked her way down the ski slope. (Goldberg 1995:205)

As seen in the sentences (147a, 147b), it is pointed out that generally, the major motion verb such as *walk* and *run* cannot participate in the *way* construction (e.g., Goldberg 1995:205), but, in contrast, in some appropriate context those verbs can actually be used, as seen in (147c). Where does the difference come from? Goldberg (1995) and Kuno et al. (2004), among others, answer to this question based on semantic and functional properties the construction is thought to have: for Goldberg (1995) the means interpretation variant of the construction must imply some “external difficulty” with which a subject referent creates a path; for Kuno et al. (2004) the construction, regardless of the interpretations, the whole sentence of the construction must describe an “unusual way” of motion. Both semantic/functional constraints can successfully exclude the unacceptable sentences (147a, 147b) but accept the sentence (147c).

For the exemplar-theoretic approach to constructions, however, such semantic/functional explanation cannot be adopted. Instead, we have to focus some *formal*, as opposed to semantic/functional, aspects of the examples. Obviously, we cannot attribute the difference in acceptability to the difference in preposition, because (147b) and (147c) share the same preposition, *down*. The next candidate is the whole prepositional phrase.

As for the two unacceptable sentences, the prepositional phrases *to the store* and *down the hallway* are considered to be somewhat frequent combinations, and actually we can find 1260 examples and 874 examples on COCA, respectively. As an approximation to the behavior of the phrases as surface patterns of some constructions, verbs appearing before the phrases within two-word gap are searched and the following results are obtained: of the 1260 examples *to the store* follows the verb *go* 393 times, accounting for about one-third of the total hits; for *down the hallway* 154 out of 874 examples contain the verb *walk*. The results show that the phrases have strong tendency to a construction which encodes the simple, probably intransitive, motion, which is perhaps the *intransitive motion* construction as seen in the following:

- (148) a. Joe walked (a mile) to the store.
 b. Bill walked (slowly) down the hallway.

The prepositional phrase in (147c), *down the ... slope* (but not *down the ski slope*, because of its rarity), in contrast, behaves differently. On COCA we can find 204 examples of the phrase and of them 20 examples accompany the verb *slide* within the two-word gap before the phrase. Though both the absolute and relative frequency is not so high, comparing other verbs appearing in the same context, the figure is considered to be significantly high: *go* occurs 9 times, which is the second-most, and *walk* occurs 8 times. This result strongly suggests that the sequence [... down the ... slope] can work as a surface pattern of a certain construction different from the simple intransitive motion construction.

Note that each of two unacceptable sentences, (147a) and (147b), is considered to contain another sequence which functions as an evoker of the intransitive motion construction, namely [*walk* ... to ...] and [*walk* ... down ...]. If those sequences do evoke the intransitive motion construction, the sentences have two evokers of it. This *dual-evoker* condition is highly likely to be the culprit of the unacceptability.

Dual-evoking would cause us to recognize the whole sentence is an instance of the dually-evoked construction, because it is natural to think that the more the pieces of evidence is, the stronger the evidence is. Those sentences, however, also include the sequence [... POSS way PREPOSITION], which evokes the *way* construction, and the construction has a *larger* meaning than the intransitive motion construction, in the sense that the former entails both the creation of path and the motion through the path but the latter only describes the motion. In this case it is usual to use the former construction because the former can do the job of the latter, but not vice versa. This is the situation in which the construction competition occurs, which probably leads to the unacceptability.

Incidentally, this construction competition-based account can be seen as saying almost the same thing as the semantic/functional account by Goldberg (1995) and Kuno et al. (2004), but from a different perspective. Their explanations are, in short, that the *way* construction should denote some *irregular* or *unusual* way of motion. If the semantics/function-based account is paraphrased in EBCG's term, the construction competition emerges: the *way* construction may compete with the simple intransitive motion construction in a certain situation, because the former describes more than the latter.

Here it should be pointed out that the acceptable sentence in (147c) is considered to be an

instance of the *manner* interpretation of the *way* construction, though Goldberg (1995:205) does not categorize in that way. Clearly, in that sentence there is no implication of the creation of path, because the path though which the subject referent moves, the *ski slope*, should be there in advance. The verb *walk* only describes the *manner* of motion, which is unusual in the situation described by the sentence. This suggests that the manner interpretation of the *way* construction is in some way related to the construction competition, which results in the *cancellation* of the creation of path meaning.

Visual representations of construction competition

The concept of construction competition seems somewhat elusive, so it is beneficial to visually illustrate the phenomenon in some way. For this purpose the description method called the *Pattern Matching Analysis* (PMA) is considered to be useful. PMA is devised by Kuroda (2000) to describe syntactic structures of sentences in a cognitively realistic way, without depending on tree structures. The detail of the method is not presented, but, simply, for a sentence it 1) provides a table called the *decomposition table* which is, regularly, an $n \times n$ matrix where n equals the number of segments the sentence contains and 2) specifies a syntactic pattern such as *S V O* for each row of the table called a *subpattern* (Kuroda 2000:24-25). The idea underlying PMA is highly similar to the exemplar theory of constructions in the sense that PMA tries to represent a segment-wise association of a certain grammatical structure, or a construction, and the process of integration of them. A sample of PMA is presented in Table 4.20, whose 0-th row provides the target sentence and each of i -th row, $1 \leq i \leq 4$, represents a subpattern associated with the i -th segment.

Table 4.20: PMA for *Ann asked (him) the way* (Kuroda 2000:24)

0.	<i>Ann</i>	<i>asked</i>	<i>(him)</i>	<i>the way</i>
1.	<i>Ann</i>	<i>V</i>		<i>(O)</i>
2.	<i>S</i>	<i>asked</i>	<i>(X)</i>	<i>O</i>
3.	<i>S</i>	<i>V</i>	<i>(him)</i>	<i>O</i>
4.	<i>S</i>	<i>V</i>	<i>(X)</i>	<i>the way</i>

Since PMA is originally devised to represent sentence-internal structures, here, for the purpose of describing constructional characters of sentences, a little modification is added to the original PMA. First the rows of a decomposition table are changed to represent a *surface pattern* with pseudo-words/phrases filling the slots of the pattern. The pseudo-words/phrases

are used for a summated set of some parts of associated exemplars, which is categorized into abstract categories (e.g., PLACE, PATH, and HUMAN), written in all capitals, or a group of words one of whose member is remarkable in terms of its frequency, represented as the remarkable member in italics with asterisk after the word (e.g., *make**). Second, for each surface pattern the name of construction which is considered to be evoked by the pattern is added on the right side of it. The process of evoking is represented as a right arrow (\Rightarrow). Just like the original PMA, the segments of a target sentence is written in bold. The modified version of PMA is named the *Pattern Matching Analysis Extended*, PMAX for short.

In Table 4.21 and 4.22 sentences with the verb *walk* slightly modified from those analyzed above are described with PMAX. For comparison, in Table ?? a PMAX for a regular example of the *way* construction with means interpretation is provided. The PMAXs illustrate the condition presented above, that is, the *dual-evoking* of the intransitive motion construction competes with the realization of the *way* construction.

Table 4.21: PMAX of **He walked his way to the store.*

0.	*He	walked	his way	to	the store .	
1.	He	<i>make*</i>	his way	to	PLACE .	\Rightarrow the <i>way</i> construction
2.	He	walked		to	PLACE .	\Rightarrow the intransitive motion construction
3.		<i>go*</i>		to	the store	\Rightarrow the intransitive motion construction

Table 4.22: PMAX of *He walked his way down the slope.*

0.	He	walked	his way	down	the slope .	
1.	He	<i>make*</i>	his way	down	PATH .	\Rightarrow the <i>way</i> construction
2.	He	walked		down	PATH .	\Rightarrow the intransitive motion construction
3.		<i>slide*</i>		down	the slope	\Rightarrow the slide-down construction

Table 4.23: PMAX of *She fought her way into the room.*

0.	She	fought	her way	into	the room .	
1.	She	<i>make*</i>	her way	into	PLACE .	\Rightarrow the <i>way</i> construction
2.	She	fought	(SOMEONE) .			\Rightarrow the action construction?
3.		<i>come*</i>		into	the room	\Rightarrow the intransitive motion construction

4.6 Concluding remarks

As have been discussed so far, the four major English constructions analyzed in a number of previous studies can be described under the current exemplar-based framework, EBCG. This section provides the brief summary of this chapter and then some remarks on the exemplar-based characterization of grammatical constructions which are obtained from the findings of the analyses presented in this chapter.

4.6.1 Brief summary

In this chapter the four major English constructions, the *ditransitive*, *resultative*, *caused-motion*, and *way* construction, are investigated under the exemplar-based framework of grammatical constructions. The methodology taken here is what can be called *surface-pattern-based* research, in which each construction is tied to describe via continuous or discontinuous sequences of words. All the constructions are analyzed through the following procedure:

- (149) a. based on previous studies or pilot researches candidates for surface patterns are specified;
- b. using those candidate sequences corpus data is researched and whether the sequences can successfully lead us to the target construction;
- c. findings from the results or further analyses of then are provided, often with some remaining problems.

Generally, it can be said that for all the four constructions typical examples are clearly and systematically described based on surface patterns. However, inversely, analyses of atypical examples cannot be said to be successful and exhaustive. For example, in the case of the ditransitive constructions examples successfully analyzed are those with pronominal indirect object whose main verb is one of the famous verbs frequently used in the construction such as *give* and *tell* and/or whose indirect object is headed by nouns strongly associated with any of those verbs; as for the *way* construction what is mainly analyzed is the *means* interpretation variant of it and the other variant with *manner* interpretation cannot be successfully analyzed.

Still, looking from a kind of meta-point of view, the descriptions, analyses and discussions provided in this chapter can be successful in the following sense. The arguments in this chapter clearly show how the exemplar-based framework works (and does not work) on describing and analyzing grammatical constructions. Because of the novelty the current frame-

work needs reasonable amount of demonstration to show what exactly it is, which we think is successful. This will probably lead to empirical verification of the framework based on the very methodology adopted here and some other critical data, resulting in the testability and falsifiability of the theory.

4.6.2 Further remarks

Based on the findings presented in this chapter, a few generalizations on the behavior of surface patterns can be drawn, and, moreover, a number of quantitative data used in the analyses of the four constructions can be reinterpreted to be reflect some general properties of sentence grammar, or at least what is claimed to be matters on sentence grammar.

Function words as distributed cues to a construction

For any of the four constructions, most of the surface patterns include *pronouns*, *articles* or *prepositions*. For example the surface pattern of the ditransitive construction include a pronoun and an articles, and some of the resultative construction contain a pronoun (e.g., [... PRONOUN mad]) or a preposition (e.g., [... into pieces]). They are, roughly speaking, members of *function words* or *closed-class words*, which are considered to be highly important jobs as to the specification of a sentence grammatical structure. For example, the pronoun sequence [I ... you] almost unambiguously lead us to the simple transitive structure, often expressed as [S V O], as exemplified in *I love you* and *I know you*. Articles in many cases tell us that the following word is a noun, or even in cases where this is not the case (e.g., when an adjective, in stead of a noun, follows, as in *the small city*), they still often function as an indicator of the *phrase* boundary, in the sense that they are in most cases show that a noun phrase starts at the place they occur.

With respect to the constructional view of grammar, this *boundary indicator* function is especially important. As noted in Chapter 3, the current theory assumes that the internal structure of a sentence is based on *element alignment* in which a set of semantic components of a construction are made to correspond to the segmented elements of the construction's formal structure (3.6.1). Therefore, what indicates the boundary of segments corresponding to the semantic components is more important than a kind of information which tells us a syntactic function or a part of speech of a certain part of a sentence.

It can be concluded that if the formal aspect of a construction can be characterized as a

series of segmented elements, the function words as segment boundary indicators do great jobs for the specification of constructions based on exemplars. In other words, exemplars can be properly organized when associated with a novel example including function words at appropriate places indicating segment boundaries, resulting in evoking a proper construction effect.

Productivity and the lexical cues

It is pointed out that there are relatively limited number of verbs which can participate in the ditransitive construction (e.g., Goldberg 1995:Chapter 5), while the *way* construction allows a large variety of verbs to participate in it. Can the current exemplar-based approach explain what makes this difference, if not necessary? The answer is: Yes.

The number of different types of items participating in a construction, which is usually called the *productivity* of the construction, can be attributed to the amount of *lexical cues* its surface patterns have. For example, the surface patterns of the ditransitive construction only specifies such few and general lexical cues as [... ARTICLE PRONOUN ...]. Seen from a different perspective, it can be said that the sequence has relatively little information which successfully leads us to the proper construction effects. On the other hand, the surface pattern of the *way* construction, [... POSS way PREPOSITION ...], is rich in lexical specificity, with the possessive determiners, the noun *way* and the prepositions. This means the sequence has information rich enough to navigate us almost unambiguously to the exemplars instantiating the target construction.

From this the following general principle can be obtained: the more lexically-specific the surface patterns of a construction is, the more productive the construction would be. In fact, for example, a putative surface patter of the resultative construction examined above, [... door open], which can be said to be lexically specific to a large degree, is used with a large number of verbs including those rarely participating the construction, namely the verbs of *sound emission* such as *crack*. Corpus search on COCA provides us with some other examples of sound emission verbs with [... door open] such as *creak*, *bump* and *buzz*, as exemplified below (underline added):

- (150) a. I creaked the door open and braced myself against the wave of stale air.
(COCA:Atlantic)
- b. He bumps the door open and enters. (COCA:Mov:HotZone)

- c. They both pulled stunners, and Mike buzzed the door open. (COCA:Analog)

Chapter 5

Exemplar theory meets Zipf's law

This chapter presents a quantitative study using corpus in order to show how surface patterns work in general. It focuses on the frequency distribution of exemplar types; specifically it investigates a distribution between the token frequency of an exemplar type and the frequency rank of the type, which is called a *Zipfian distribution* (Zipf 1935, 1949). This distribution is known to represent a kind of bias in human perception and cognition.

Given the very exemplar-based nature of the theory, EBCG has some difficulty in accounting for the whole picture of a language. It is even harder for EBCG to empirically demonstrate the pattern-based nature of construction recognition using corpus data. Consequently it is clear that EBCG needs some complementary measures which can cover the sparseness of EBCG's (potential) findings.

A corpus analysis presented in this chapter is, therefore, done to supplement the sparseness. We, as researchers, cannot fully know the whole picture of a linguistic phenomenon related to human cognition due to the lack of data availability and, more crucially, the unknowability of human mind. We never know how much and what kind of input a specific individual has obtained so far; we never even know how the input stimuli are processed in his/her mind.

5.1 Zipfian distribution of rank-frequency correlation

It is well known that quite a large number of linguistic phenomena show Zipfian distribution in the correlation between their type ranks and token frequencies. Zipfian distribution is a distribution which obeys what is called *Zipf's law*, characterized as a kind of *power law*. Specifically, the Zipfian distribution between type rank and token frequency, abbreviated as

Zipfian rank-freq distribution hereafter, is an inverted correlation between a rank of a type and the frequency of the type obeying a power law. Power law distribution becomes linear if we take the *logarithms* of the two variables, in this case the rank and the frequency.

A sample of the distribution is illustrated in Figure 5.1 and 5.2, the latter of which is a log-scaled version of the former. They are the frequency distribution of words in a well-known mid-scaled balanced corpus of American English, *Brown Corpus*.¹⁾

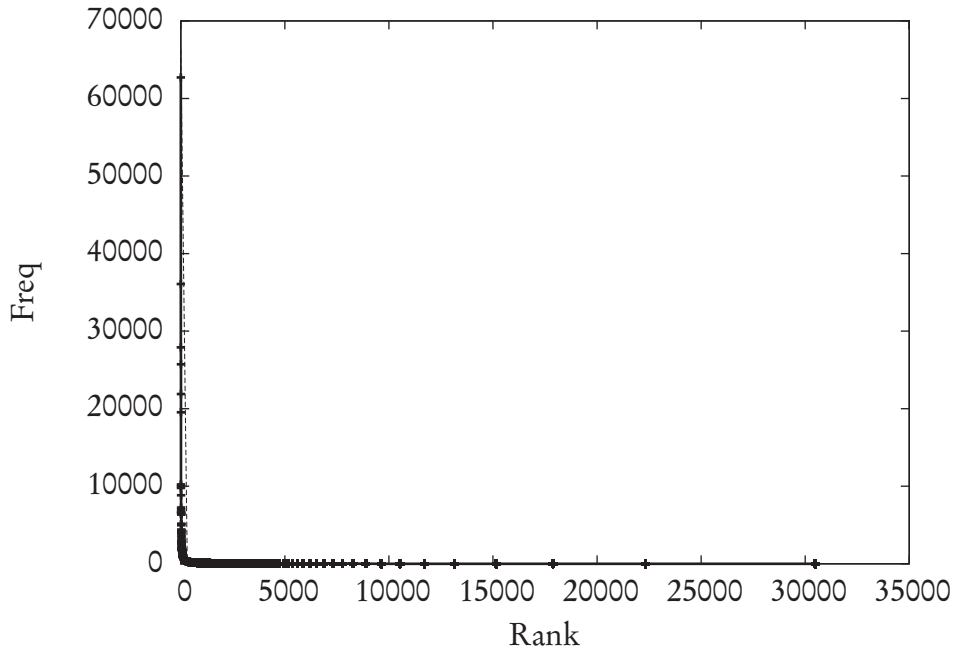


Figure 5.1: A rank-freq distribution of *Brown Corpus*

The Zipfian rank-freq distribution can be characterized by the invariant proportion of the frequency in each rank of words to that of the topmost word. Letting S_1 be the frequency of the topmost word, the frequency of a rank R , S_R , can be given by the following formula:

$$S_R = \frac{S_1}{R} \quad (5.1)$$

For example the most frequent word in *Corpus of Contemporary American English* (COCA, Davies 2008-), *the*, occurs 25064010 times and the second most word *and* occurs 12348424 times; if these values are substituted into the above formula, the following value is obtained:

$$S_2 = \frac{25064010}{2} = 12532005$$

¹⁾The data was retrieved from *Natural Language Toolkit* (NLTK, Bird et al. 2009), a library of a script language *Python* (<http://python.org/>), which provides a variety of language processing tools and database.

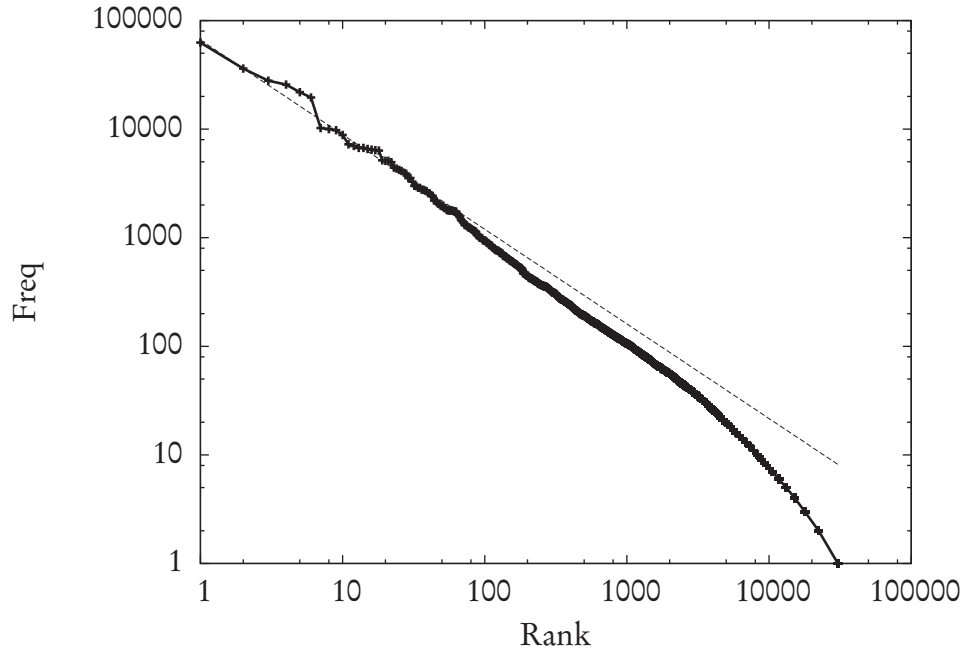


Figure 5.2: A rank-freq distribution of *Brown Corpus* (log-scaled)

The obtained value is almost the same as the observed frequency of the second most word, 12348424.

5.1.1 Cognitive reality of Zipf’s law

Zipfian distribution may not only be a mere tendency found in frequency distribution, but reflect some property of our numerical perception or cognition. If we often find frequency distributions obeying Zipf’s law, there should be a motivation or simply a reason for the frequent formation of the distribution. One leading candidate for the motivations is the nature of our quantity perception.

Traditionally the logarithmic nature of human perception is modeled via what is called *Weber-Fechner law*. It states that perceived intensity P can be approximated by the following formula:

$$P = K \log \frac{S}{S_0}$$

where K is a parameter estimated via empirical data, S_0 is a minimal threshold of perception and S is physical stimulus intensity (Varshney & Sun 2013:30).

Correlation between perceptual intensity and stimulus intensity obeying Weber-Fechner law is illustrated in Figure 5.3 and 5.4, in the latter of which x-axis is log-scaled. Those graphs clearly show that in order for us to double our perceptual intensity P (plotted in y-axis), we need to *square* the stimulus intensity S (plotted in x-axis). As for the frequency of words in a certain context, therefore, logarithmic, or exponential, frequency distribution can be seen optimal to us; we are good at detecting the difference in frequency between some word and another when the difference is logarithmic/exponential.

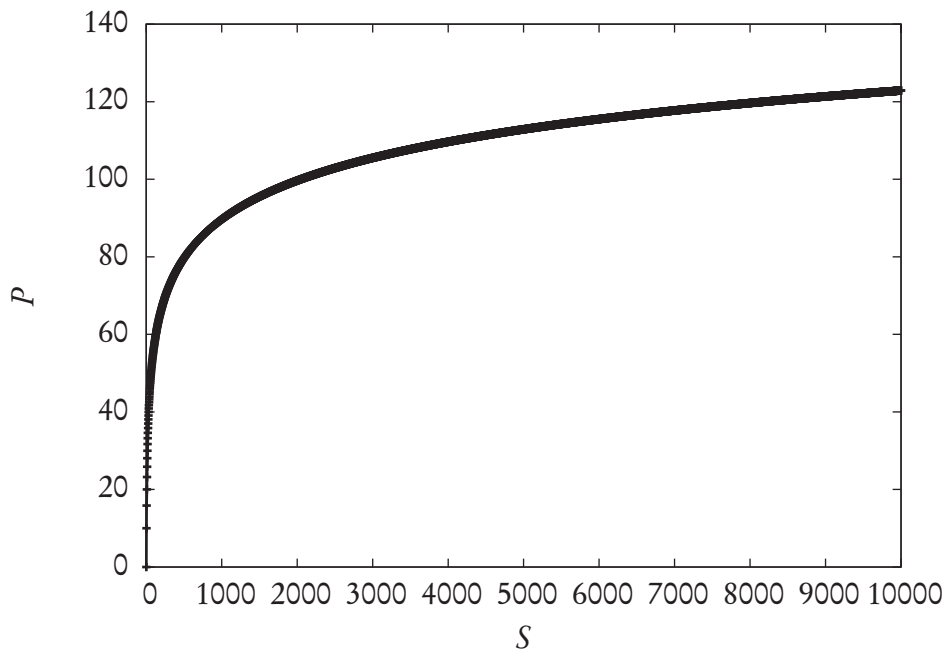


Figure 5.3: Logarithmic perception of quantity [where $k = 10$ and $S_0 = 2$]

From this it follows that the idealistic frequency distribution for us humans is discrete one whose elements distributed with a log interval scale. That kind of distribution obviously is Zipfian. In this connection, Varshney & Sun (2013) investigates the logarithmic nature of human perception and cognition of quantity. They explain it through evolution, the statistics of nature and error reduction. Especially, they argue that the logarithmic nature of our quantity perception would be beneficial in terms of *error minimization*. (Sun et al. 2012:see for the detail of their model for error minimization). Suppose our quantity perception is linear, not logarithmic, and the frequency distributions found in nature is logarithmic. In such cases, for example, encountering a certain phenomenon eight times and another phenomenon sixteen times are perceived in the way in which the latter happens twice as frequently as the former; if we encounter the former ten times the perceived difference between the former and the latter may differ to some degree. If logarithmic, the difference is much smaller: $\log 8$

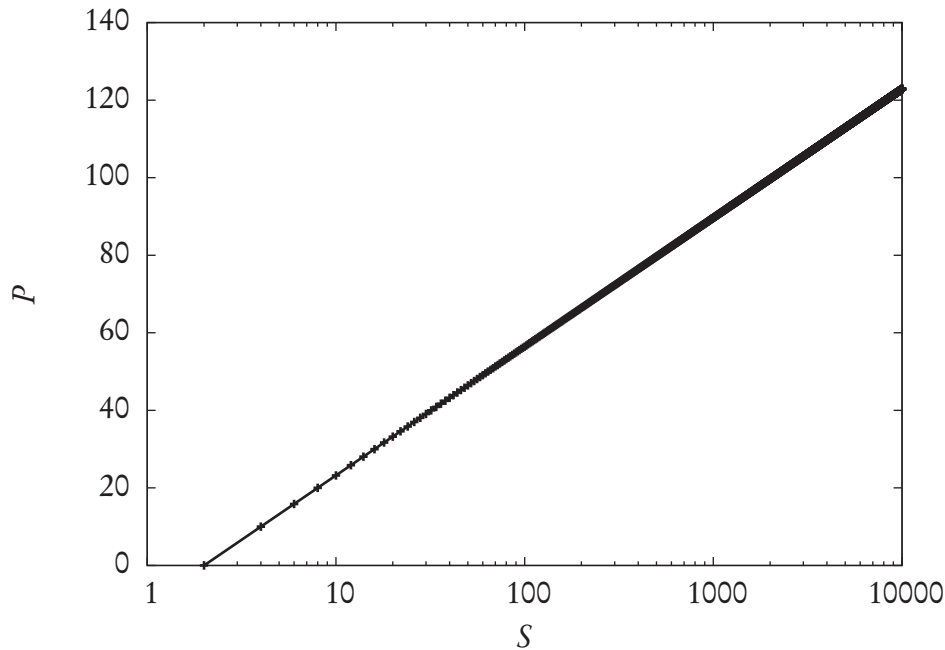


Figure 5.4: Logarithmic perception of quantity with X-axis log-scaled

and $\log 10$ with the base of two equals just 3 and about 3.32, respectively. Any phenomenon may occur irregularly frequently in some occasions and in that case our perception, if linear, would be misguided by the irregular frequency. In contrast, logarithmic perception of quantity can minimize the effect of those noisy irregularity, if the perceived quantity itself is also distributed logarithmically.

5.1.2 Relation to language acquisition

Goldberg (2006:75-83) points out that for infant language learners an argument structure construction is most learnable when the frequency distribution of verbs of it is *skewed*, which means the frequency of the topmost type is extremely high while others are less than half as frequent as the topmost one. This kind of skewed frequency distribution clearly shows Zipfian nature (Goldberg 2006:76).

Specifically, she and her colleagues (Goldberg et al. 2004) researched a corpus of children-mother dialog (Bates et al. 1988) from CHILDES database (MacWhinney 2000) and found that from mothers' speech the most frequent verbs in some arguments structure constructions account for quite large proportion of the overall construction (see Table 5.1). This means that the input of constructions children gain shows skewed nature in its frequency distribution.

They also conducted an experiment to investigate the effect of skewed input on language

Table 5.1: The most frequent verbs in three constructions from mothers' speech of Bates et al. (1988) (adapted from Table 4.2 in Goldberg (2006:76) and data given in Goldberg et al. (2004:296))

Constructions	The topmost verb	#verb types	Example
Subj V Obl	<i>go</i> (39% [136/353])	39 verbs	<i>I went to the store.</i>
Subj V Obj Obl	<i>put</i> (38% [99/256])	43 verbs	<i>Marty put the milk in the fridge.</i>
Sub V Obj Obj ₂	<i>give</i> (20% [11/54])	13 verbs	<i>Pat gave Chris a book.</i>

learning (Casenhiser & Goldberg 2005). They created a novel argument structure construction and tested how much easily the construction was learned if the frequency distribution of verbs in it is skewed, compared to a less skewed, balanced frequency condition. Using video clips, they showed several scenes representing a particular event with audio descriptions of the scene, which were the input stimuli of the novel construction, to fifty-one native English speaking children aged from five to seven (mean = 6.4) (Casenhiser & Goldberg 2005:502).

Both in balanced frequency condition and in skewed frequency condition the number of verb types are five. In the balanced frequency condition the verbs are distributed as follows: two verbs occurred once and the other three verbs occurred twice (1-1-2-2-2); in the skewed frequency condition one verb occurred four times and the other four verbs occurred only once (4-1-1-1-1) The total sum of token frequency was the same, namely eight (Casenhiser & Goldberg 2005:503).

After showing the videos they tested whether the children learned the construction or not, using a force-choice comprehension task. They showed two clips whose scene were fundamentally different and one of which was the correct one in the sense that it represented the intended scene as the meaning of the construction. The two clips were shown with one audio description of the novel construction in question and they "asked [the children] to point to the film clip that corresponded to the description that they heard" (Casenhiser & Goldberg 2005:503).

The result was really clear: children's performance in skewed frequency condition was significantly higher than that in balanced frequency condition (Casenhiser & Goldberg 2005:503-504). Similar results were found in the case of adults in an experiment by Goldberg et al. (2004:299-302).

Ellis & Ferreira-Junior (2009) also argue about the relation between Zipfian rank-freq distribution and learning of constructions. They focus on adult second language learning, but

found quite similar results based on corpus analysis. They researched conversational data between English native speaker interviewers and non-native speaker interviewees provided from the European Science Foundation (ESF) project (Dietrich et al. 1995; Feldweg 1991; Perdue 1993). The result of their analysis was that the frequency distribution of verbs in several argument structure construction they investigated shows Zipfian nature both in native speaker interviewers' speech and in non-native speakers' (Ellis & Ferreira-Junior 2009:199-200).

5.1.3 Some notes on the mathematics of Zipf's law

It may deviate the main story and purpose of this chapter in particular, and even this dissertation in general, but here some notes on the *mathematics* of Zipf's law should be provided, because this helps understanding of the results obtained from the quantitative studies given below.

The equation (5.2) is the regression of the Zipfian rank-size distribution, where R is the *rank* and S is the *size*, or *frequency*. This equation clearly shows that the two variables, R and S , are inversely proportional if log scaled. Compare with an equation of simple inverse proportional relation, $y = -\beta x + \alpha$, substituting $\log R$ with y and $\log S$ with x .

$$\log R = \alpha - \beta \log S \quad (5.2)$$

This regression can be expanded as follows:

$$\begin{aligned} \log R &= \alpha - \log[S^\beta] \\ \exp[\log R] &= \exp[\alpha - \log[S^\beta]] \\ R &= \frac{\exp \alpha}{\exp[\log[S^\beta]]} = \frac{e^\alpha}{S^\beta} = e^\alpha S^{-\beta} \end{aligned} \quad (5.3)$$

Here, if we let a be e^α and b be $-\beta$, it can be rewritten as follows:

$$R = aS^b \quad (5.4)$$

This is practically the same as an equation of *power function*:

$$y = ax^b \quad (5.5)$$

where the two variables x and y correspond to S and R in the above equation (5.2), respectively.

It is known that in the Zipf's law the parameter β is equal to one, hence b being -1 . From this it follows that if we fit a distribution obeying Zipf's law using power law function (5.5), the parameter b should be estimated as -1 .

As for the parameter a , the equation (5.1) should be taken into count. The equation is provided again for convenience:

$$S_R = \frac{S_1}{R} \quad (5.1)$$

Obviously this equation can be transformed into the following:

$$R = \frac{S_1}{S_R} = S_1 S_R^{-1} \quad (5.6)$$

Further, in this equation S_R represents a specific value of the frequency variable at the rank R , and therefore it can also be expressed as generalized form S , which is identical to the variable represented by the same letter S in the regression (5.2) and its transformed version (5.4). Consequently, it can be converted into the following, substituting S_R with S :

$$R = S_1 S^{-1} \quad (5.7)$$

Now, comparing (5.7) with the transformed version of Zipf's law regression, (5.4), it becomes clear that the constant $a' (= e^a)$ is the same as S_1 . In summary, the frequency distribution obeying Zipf's law should fit a power function $y = ax^b$ where $a = S_1$ and $b = -1$.

Further, this equation can be transformed into a function in terms of the variable x :

$$\begin{aligned} y &= ax^b & (= 5.5) \\ x^b &= \frac{y}{a} \\ x &= \left(\frac{y}{a} \right)^{1/b} \end{aligned} \quad (5.8)$$

Here it should be noted that by convention, a plot of a frequency distribution is usually done

in the way in which the variable x corresponds to a rank of word types and y corresponds to a frequency of words at each rank. This means the transformed version of the power function, $x = (y/a)^{1/b}$, is, in an actual plotting, equal to $y' = (x'/a)^{1/b}$, where $y' = x$ = the frequency variable and $x' = y$ = the rank variable. In addition, in this transformed version the value of b should be estimated as -1 if the distribution truly obeys Zipf's law, because the right side, $(x'/a)^{1/b}$, is the same as the right side of the equation (5.1) if the exponent b is equal to minus one the two equation become equal:

$$y' = \left(\frac{x'}{a}\right)^{1/-1} = \left(\frac{x'}{a}\right)^{-1} = \frac{a}{x'} \quad (5.9)$$

Furthermore, let b' be $-1/b$, the negative version of the exponent $1/b$. Then the above transformed equation can further be transformed as follows:

$$y' = \left(\frac{x'}{a}\right)^{1/b} = (a^{-1}x')^{-b'} = a^{b'}x'^{-b'} \quad (\text{where } b' = -1/b) \quad (5.10)$$

Finally, let a' be the complex parameter $a^{b'}$:

$$y' = a'x'^{-b'} \quad (5.11)$$

This looks more like the original form of the power function $y = ax^b$. In this final transformation the parameter b' should be 1 if the distribution in question truly obeys Zipf's law, because when the value of the original b , b' is equal to 1: $-(1/-1) = 1$.

Now, since all the elements in the transformed equation in (5.11) are marked with prime symbol, they are replaced by the simple capitals, A , B , X and Y :

$$Y = AX^{-B}, \quad \text{where } A = a^{b'}, B = -1/b, X = y, Y = x \quad (5.12)$$

Table 5.2 provides a brief summary of several equations related to the regression of Zipfian distribution in terms of the variables and parameters.

Table 5.2: Parameters and variables in various equations

	$\log R = \alpha - \beta \log S$	$y = ax^b$	$Y = AX^B$
rank	R	y	X
freq	S	x	Y
parameter 1	α	$a(= e^\alpha)$	$A(= a^{-1/b})$
parameter 2	β	$b(= -\beta)$	$B(= -1/b)$

5.1.4 How Zipfian should a distribution be?

Given that the argument here is about human perception and/or cognition of linguistic input, the mathematical characterization of the distribution may not be something we should pursue. It is, therefore, necessary to extract some essence of the law important to our perception and/or cognition of linguistic input, and, if possible, we should connect the extracted to some mathematical properties in order to evaluate the distribution in a somewhat rigid way.

Here let us remember that Goldberg et al. (2004) and Casenhiser & Goldberg (2005) show the importance of *skewed* input. Their experiments employed sequences whose topmost types were extremely frequent compared to others. This means that the input distribution can be optimal to learners if the difference in frequency between the topmost item and the others is large, whether the whole shape of the frequency distribution obeys Zipf's law or not. Put otherwise, even if the overall shape shows Zipfian nature, the distribution cannot be said as optimal without the topmost item being by far the most frequent.

At the same time, however, in view of the logarithmic nature of our perception mentioned above, the difference in frequency between items of each rank should be logarithmically separated, because, if not, we cannot perceive the difference of frequency and hence become unable to learn the context in which the words in question distribute as a pattern, or a construction.

The question here is: how can we know the distribution in question is optimal to language learners? In this respect the mathematics of Zipf's law or power law in general gives us some hints. Let us look at the figure (5.5), which shows plots of an equation $Y = AX^{-B}$ where A is fixed as 1000 and B is varied: $B = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2\}$. Seeing what happens when the value of B changes, it becomes clear that the parameter B represents, as it were, a degree of difference between the topmost item and others. In fact, the parameter B represents a *slope* of a line if log scaled.

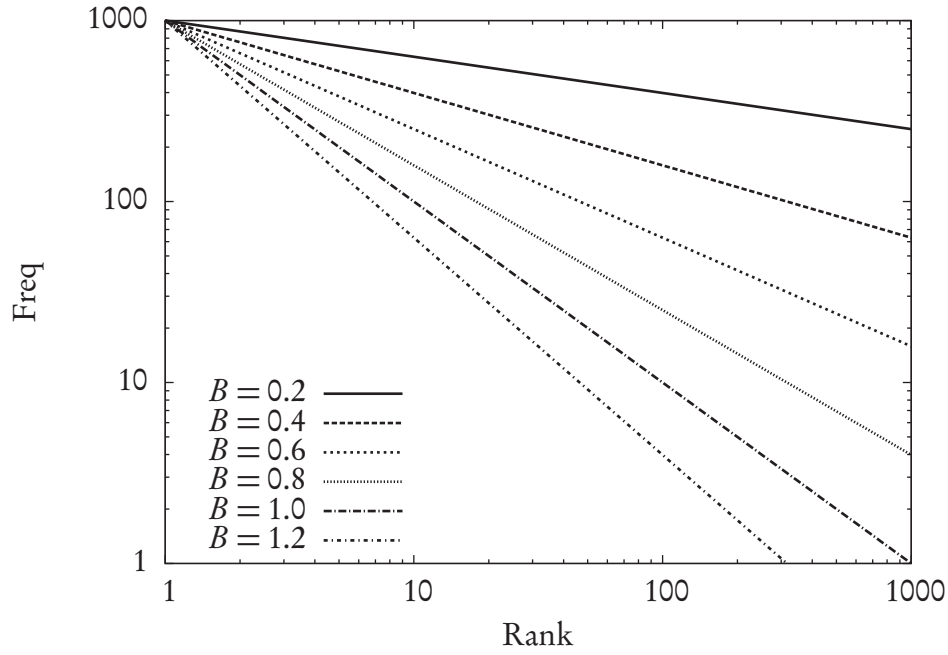


Figure 5.5: How the graph changes according to B

Consequently, we may be able to conclude that if we fit the distribution to a transformed power law function $Y = AX^{-B}$, the estimated value of the parameter B is a good indicator of the optimality of the distribution as a linguistic input, even though it may somewhat deviates Zipfian distribution. In other words, the greater the value of B is, the more optimal the distribution is.²⁾

However, in addition to B , the role of the other parameter, A , should also be taken into consideration. As seen before, in rank-freq distributions A corresponds to the frequency of the topmost item, S_1 , if B is equal to 1. Yet it is unclear whether this also holds true for the cases in which B is not equal to one, so it is necessary for us to see how the value of A changes when the value B is not equal to 1.

In order to know the nature of the parameter A , let us do some calculation. Now we can easily see the nature of the parameter A by looking at the case where $X = 1$. The frequency of the topmost word, Y_1 , is calculated as follows:

$$Y_1 = A \times 1^B = A \quad (5.13)$$

²⁾Note that the parameter B is different from the original one, b , but, though their scales are different, the magnitude relationship between b and B is the same. For example, compare b s and B s when $b = 1$ and $b = 1.2$. For convenience let B_1 be the value of B when $b = 1$ and B_2 be the value of B when $b = 1.2$. Now B_1 and B_2 are calculated as follows: $B_1 = -(1/1) = -1$; $B_2 = -(1/1.2) = -0.8333 \dots$. Clearly, B_2 is greater than B_1 , in the same way as the relationship between the b values, namely that between 1.2 and 1.

This tells us that the value of A equals the frequency of the topmost word as in the value of the original parameter a , whatever value B is.

Then let us calculate the frequency of the second most word in a distribution whose topmost item occurs 1000 times. If the value of B is 0.8, the following value is obtained:

$$y'_2 = A \times 2^{-0.8} = y'_1 \times 2^{-0.8} = 1000 \times 2^{-0.8} \approx 574.35$$

The obtained value, which is about 574.35, surely is greater than the expected value of Y_2 when $B = 1$, namely 500. Therefore the difference in frequency between the topmost and the second most word becomes smaller comparing to the true Zipfian distribution where $B = 1$.

In conclusion, even if we take the value of A into consideration, the value of B , or b , is still crucial, and therefore the value can be said as a good indicator of the overall shape of the distribution is optimal to our perception. In addition, given that the parameter A is estimated as the same value of the frequency of the topmost item, Y_1 , whether A value is equal to Y_1 or not is also tells something about the distribution. If the value differs greatly from Y_1 , the overall form of the distribution itself cannot be identified as obeying power law.

5.2 Analysis of large balanced corpora

Here a hypothesis is presented on the relation between plausibility to assume a sequence as a surface pattern and the nature of frequency distribution in words appearing specific position in, before and after the sequence:

- (151) If a sequence q is a surface pattern of a certain construction c , words appearing in a specific position related to (i.e., in, before or after) the sequence q show Zipfian rank-freq distribution.

In the remaining of this section several patterns will be investigated by checking whether the frequency distribution of words in a certain position of patterns shows a Zipfian nature or not. The patterns investigated are:

- (152) a. Patterns including major prepositions (to examine the general tendency)
 b. [X PRONOUN ARTICLE] (on the ditransitive construction)

- c. [X DETERMINER NOUN to death] and [X DETERMINER head off] (on the resultative construction)
- d. [X ARTICLE ball inside] (on the caused-motion construction)
- e. [X POSS way PREPOSITION] (on the *way* construction)

Before going to the main research part, however, some preliminary discussions are provided as to 1) the data employed in all the researches and 2) comparison between patterns and non-patterns.

5.2.1 The data

The data used in this study is a large-scaled balanced corpus of American English, called *Corpus of Contemporary American English* (COCA, Davies 2008-). The corpus is composed of more than 450 million English words with a variety of genres including spoken, fiction, popular magazines, newspapers, and academic texts. As of October, 2013, it contains data from 1990 to 2012; the corpus is updated annually, augmented with about 20 million words (see data summary shown in Table 5.3).

Table 5.3: Data summary of COCA

genre	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC	TOTAL
#words	95,385,672	90,344,134	95,564,706	91,680,966	91,044,778	464,020,256

All the words in the corpus are tagged with part of speech, but no information about syntax is annotated. The corpus can be searched using a web interface (<http://corpus.byu.edu/coca/>), which is provided by Mark Davies from Brigham Young University. You can search data with surface forms, lemmatized forms and part of speech tags. Tags can be used with surface forms and lemmatized forms.

This corpus is used for the following reasons:

- (153) a. It is the largest balanced corpus of contemporary English which is freely available;
- b. The interface of it allows us to search its data flexibly.

The primary reason is the first one (153a). The purpose of the current study is to examine the nature of frequency distribution, but a corpus with relatively small amount of data provides limited opportunity to investigate distributional features compared to larger ones. Generally,

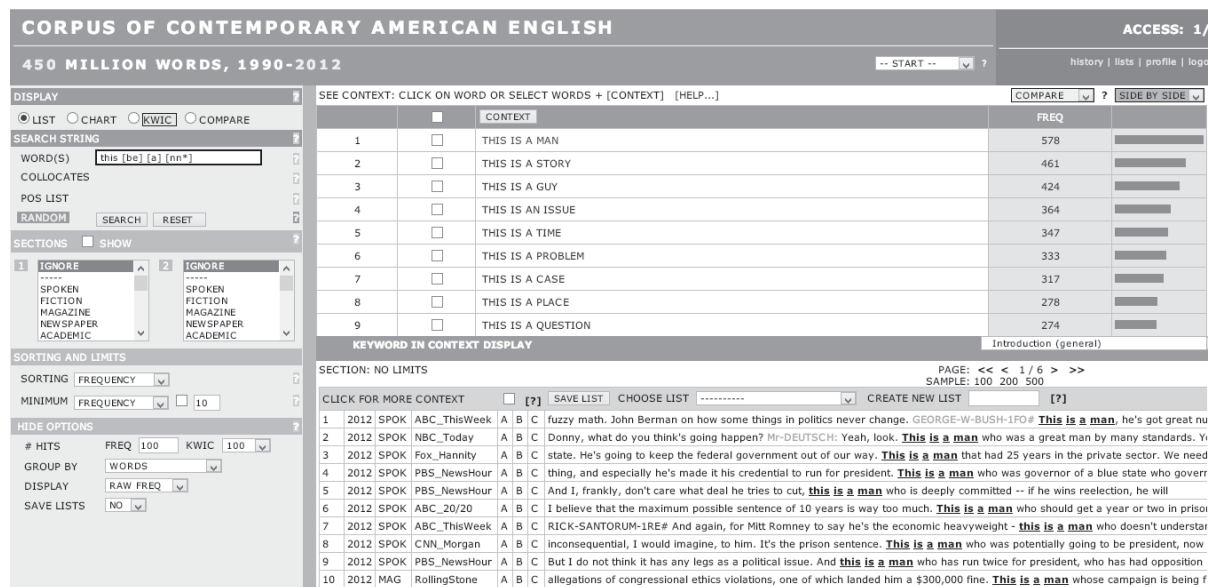


Figure 5.6: The web interface of COCA

the smaller the scale of a corpus is, the sparser the data it has is. Data sparseness, in turn, causes a difficulty in capturing distributional features of items, because in order to examine the frequency distribution of a pattern or the like, we need a certain amount of sample data exemplifying the pattern.

How the corpus scale affects the distributional nature of a pattern

For example, let us make some comparison of the behaviors of a pattern between two corpora with different scales. The pattern examined here is what can be seen as a pattern evoking the caused-motion construction, [X ARTICLE *ball inside*]. The pattern is chosen because it is quite probably involved with the caused-motion construction, as shown in the previous chapter (4.4), but it is somewhat rare in terms of frequency.

The comparison is made between COCA and a large-scaled balanced corpus of British English called *British National Corpus* (BNC). BNC contains about 10 million English words, including various kinds of genres which are almost the same as COCA. Therefore it can be said that, though the of English differs between American and British, the two large-scaled corpora share reasonably comparable features. At the same time, they differ in size: COCA is about four times as large as BNC. In consequence they are suitable to see how the difference of size affects distributional characters of words appearing in a pattern.

BNC is searched using a web interface also provided by Mark Davies, which is available

at <http://corpus.byu.edu/bnc/x.asp>, which is called *BYU-BNC* (Davies 2004-).³⁾ The result is shown in Table 5.4, which lists the words appearing the *X* slot in [X ARTICLE *ball inside*]. As for the result obtained from COCA, only the words occurring more than once are listed, and they are a part of what we already saw in Table 4.16 in the previous chapter.

Table 5.4: Comparison of *X* in [X ARTICLE *ball inside*] between COCA and BNC

COCA			BNC		
rank	X	freq.	rank	X	freq.
1	get	16	1	at	2
2	pound	9	2	turn	1
3	have	3	2	touch	1
4	pump	2	2	switch	1
4	run	2	2	square	1
4	dribble	2	2	slip	1
4	work	2		⋮	
	⋮			total	7
		total 52			

As is evident from the table, the pattern in COCA shows Zipfian nature in the sense that the topmost item, *get*, occurs extremely frequently compared to others, as we saw in 4.4. In fact the distribution obeys power law (see Figure 5.27): the estimated value of A in $Y = AX^{-B}$ (≈ 16.46) is almost equal to the observed frequency, namely 16, and that of the parameter B is larger than one (≈ 1.34).⁴⁾

On the other hand, we can obtain only a few examples of the pattern from BNC. The result cannot be fit to the power law regression because of the number of items: there are only two samples and hence no estimation can be done in terms of the shape of a distribution they are in. This result strongly suggests that relatively small scale causes a difficulty in examining distributional character of somewhat rare patterns.

This difference, however, may be reduced to the dialectic difference between American

³⁾Its appearance and functions are almost the same as COCA's, so the detailed description of it is omitted here.

⁴⁾The regression and plotting is done used with a software named *gnuplot* (<http://www.gnuplot.info/>), which is a command-line-based graph drawing utility. The version of *gnuplot* is 4.6, built as an application in the *Cygwin* system (<http://www.cygwin.com/>), which is a collection of tools providing a kind of virtual Linux environment via command line.

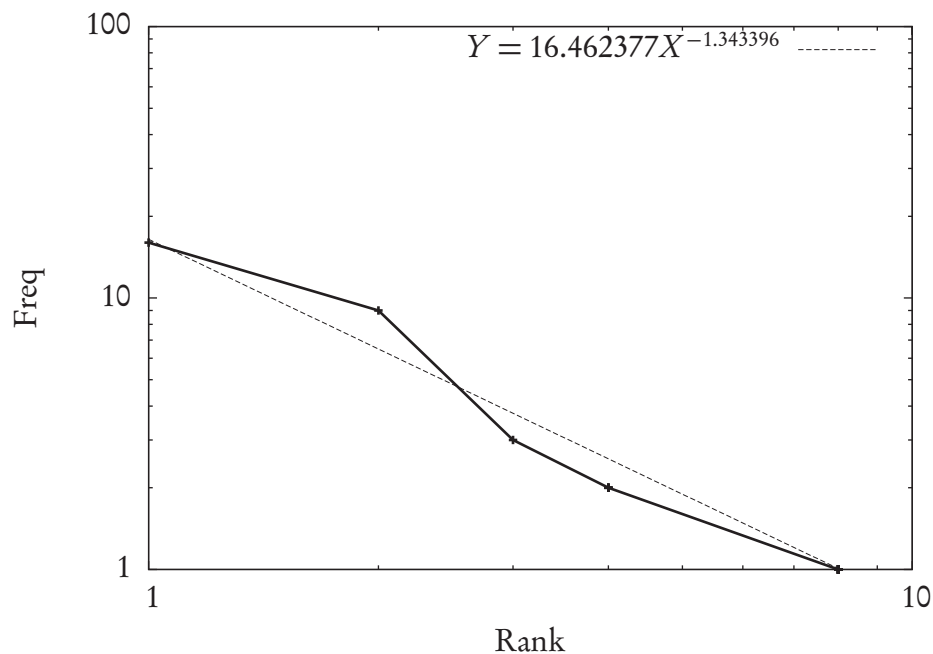


Figure 5.7: X in [X ARTICLE *ball inside*]

and British: in the first place the expression is more famous in American English than British and the result may differ in that way for that reason, not because of the difference in size.

However, even if the difference in the results presented above is caused by the dialectic gap between American and British English, it does not mean that the difference in size has no effect on observed distribution of words appearing in a specific position of a certain pattern. In this case, if we can find more examples of [X ARTICLE *ball inside*] in larger-scaled corpus of British English than BNC, it can be said that the difference is indeed caused by the size gap.

Fortunately, there is a far larger scaled corpus of English, including both British and American, called *Google Books Ngram Corpus* (Michel et al. 2011). The corpus contains about as many as hundreds of billions of words from English books published from 1950s to 2000s, which are extracted from the Google Books.⁵⁾ We can search its data and get a result visually at <http://books.google.com/ngrams>, the site of *Google Books Ngram Viewer* (see Figure 5.8).

The viewer allows us to search words or word sequences, or n-grams, with respect to genres and dialects, so we can examine dialectic differences if we search a certain word or n-gram in American and British English separately. Figure 5.9 shows the result of a search in which the sequence [ARTICLE *ball inside*] is queried. The X axis represents the year in which the sequence is used and the Y axis shows its relative frequency. From the result it is

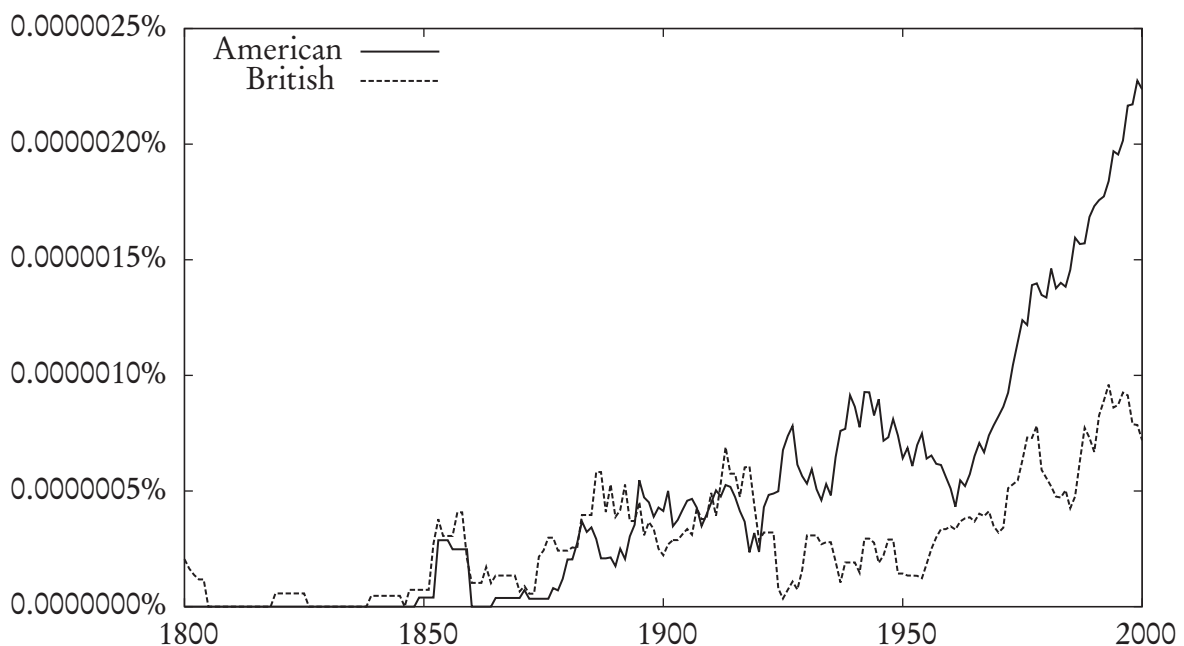
⁵⁾<http://www.google.com/intl/en/googlebooks/about/index.html>

Google books Ngram Viewer



Figure 5.8: Google Books Ngram Viewer

clear that the phrase is indeed more famous in American English than in British English; in American English it occurs more than four times as frequent as that in British English, in 2000.

Figure 5.9: [ARTICLE *ball inside*] in American and British English

However, this graph does not tell us how many of examples are actually found in each dialect. In order to know the number of occurrences, we need another interface of Google Books corpus. Fortunately again, there is another interface with which we can get the exact

number of occurrences. The interface is also provided by Mark Davies, available at <http://googlebooks.byu.edu/> (Davies 2011-). The British component of the corpus can be searched separately from American.

From the British component of the corpus which amounts to 34 billion words, we can find 173 examples of the sequence [ARTICLE *ball* inside]. This is about 25 times as many occurrences as in BNC (= 7). The number of occurrences, 173, is not so small and actually exceeds that found COCA, namely 52, to a great degree. In consequence, it can be concluded that the difference in corpus size actually matters.

The flexibility of interface

The second reason for using COCA, (153b), is related to the interface of Google Books Ngram Corpus. There is an interface for it which is, as mentioned above, also provided by Mark Davies and hence is similar to that of COCA, but there are some crucial differences between them. In terms of the scale of corpus, Google Books Ngram Corpus is indeed much better than COCA because the former has about 75 times as many words as the latter (34 billion vs. 450 million) and therefore the latter is suitable for the investigation of distributional characters of patterns, but, in terms of its interface, COCA is far better than Google Books Ngram Corpus. For example, for now we can search collocates neither with part-of-speech tags nor lemmas, but only with single words (so we cannot find any items occurring just before [PRONOUN ARTICLE], which are considered to be verbs in the ditransitive construction). In addition, KWIC search function has not been provided so far. Contexts in which retrieved words or sequences are used can be accessed, but the contexts are only provided as a result of *Google Search*, which largely depends on its search algorithm and hence sometimes provides unexpected results.⁶⁾

⁶⁾For example, results of searching sequence [*kick* PRONOUN ARTICLE] include examples with punctuation marks such as a period and a question mark intervening the sequence, as exemplified below (underline added):

- (iv) The blow that had connected was harder than that from any horse that had ever kicked him. A low gasp of surprise rose from those who watched.
(<https://books.google.co.jp/books?id=iEYctq1VLAYC&pg=PA289&dq=%22kicked+him+a%22>)

5.2.2 Comparison between patterns and non-patterns

In order to verify the above raised hypothesis, it is necessary to show not only that a sequence which can be seen as a surface pattern displays a Zipfian nature, but also that sequences which cannot be regarded as patterns do not involve Zipfian distribution.

For this purpose, first, we compare two sequences one of which composes a grammatical phrase and the other of which does not. The sequences examined are:

- (154) a. [*the book* X]
 b. [*book the* X]

If the hypothesis (151) is right, the sequence (154a), which composes a noun phrase and hence is assumed to evoke some construction effect, shows its Zipfian nature in frequency distribution of words appearing in the *X* slot, while for the latter sequence (154b) words in the slot are not distributed in a Zipfian way.

The two sequences are searched using COCA. The result are showed in Table 5.5 and 5.6 and Figure 5.10 and 5.11. Their difference is quite apparent: as seen in the Figure 5.10 the overall shape of the distribution can be seen to largely fit to power law, the value of *B* is a little larger than one (≈ 1.060) and *A*'s value (≈ 3569.365) is almost equal to the frequency of the topmost word *be*, according to Table 5.5; on the other hand, Figure 5.11 shows that, though the overall shape nicely fits to power function and indeed the value of *A* is almost precisely estimated as seen in Table 5.6, the value of *B* is really low, about as half as one (≈ 0.509).

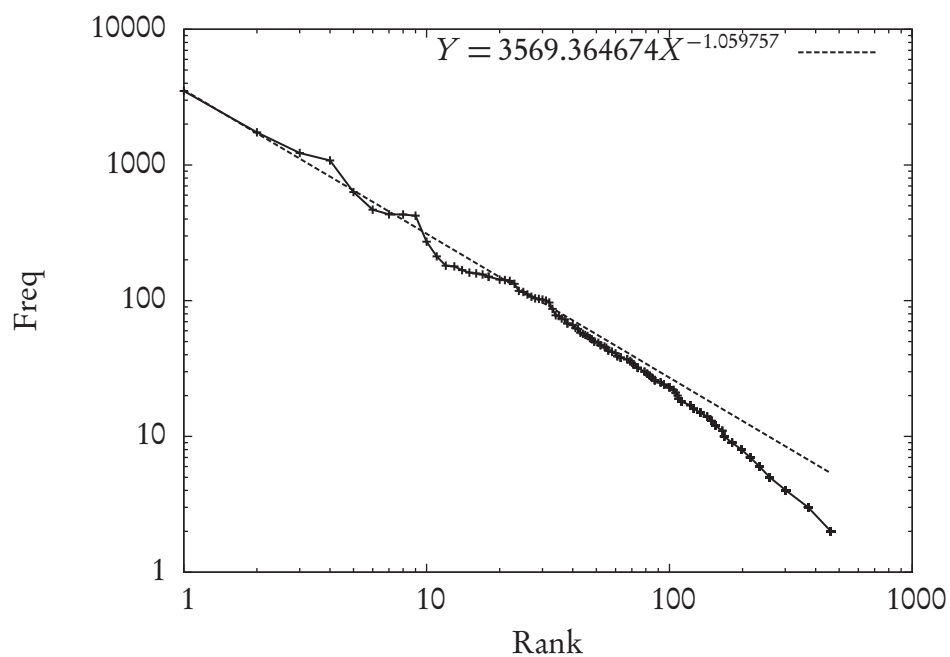
This result strongly supports the hypothesis (151). It provides us with a good criterion to separate patterns and non-patterns: even if words in a specific position of a sequence are distributed in a power law fashion, the sequence cannot be seen as a surface pattern when the parameter *B* is far smaller than one, perhaps about a half as small as one. Of course this is not a strict criterion by which a sequence is judged in terms of whether it is a pattern or not, but the degree, about a half of one, can be seen as a good guide.

5.2.3 Patterns including major prepositions

Now let us compare patterns with some prepositions. Some of the English prepositions are used in various syntactic context and are highly polysemous. This character makes them less informative as to the recognition of constructions; in other words, some polysemous prepositions rarely act as evokers of constructions.

Table 5.5: X in [*the book X*]

rank	X	freq.
1	be	3505
2	of	1741
3	's	1226
4	and	1079
5	that	631
6	on	468
7	have	433
8	to	432
9	in	423
10	as	272
	\vdots	
total		29652

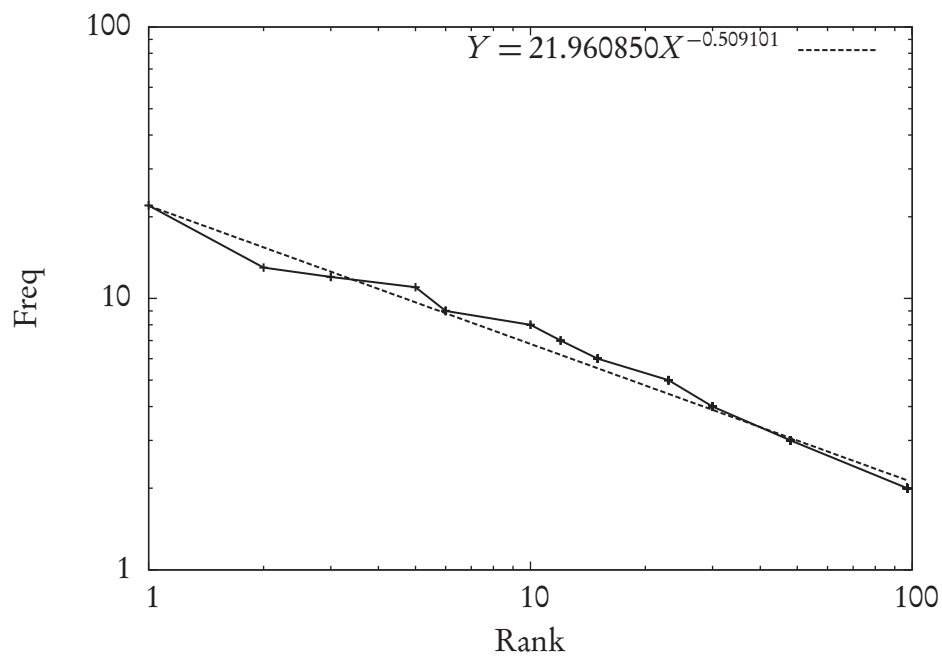
Figure 5.10: X in [*the book X*]

Preliminaries

The typical example of such polysemous prepositions is *in*. The preposition *in* hardly evokes any construction effects by itself. In contrast, other prepositions such as *to* and *into* can be thought to evoke some construction effects even by themselves, such as the intransitive

Table 5.6: X in [*book the X*]

rank	X	freq.
1	new	22
2	end	13
3	american	12
3	art	12
5	great	11
6	other	9
6	power	9
6	same	9
6	last	9
10	first	8
10	way	8
	⋮	
total		1295

Figure 5.11: X in [*book the X*]

motion construction (155a), the caused-motion construction (155b) and some noun phrases (155c):

(155) a. I went to the store.

b. She took me to the station.

c. an investigation into the nature of language

In this connection, it is beneficial to see how differently the three prepositions behave in the same context. The context is [*I ... PREPOSITION the room*]. The preposition *in* can be used with several types of verbs in the context:

- (156) a. I stood in the room.
 b. I danced in the room.
 c. I walked in the room.
 d. I went in the room.

However, those sentences somewhat differ in their meaning. The first sentence, (156a) exemplifies what is called the *existential construction*, meaning the subject, in this case *I*, exists in the place specified by the oblique, whereas the forth sentence instantiates the *intransitive motion* construction. The second sentence (156b) instantiates a construction similar to the existential construction but somewhat different from it, which can be called the *bodily action* construction or the like,⁷⁾ meaning the subject does some bodily action. The prepositional phrase describes the place the action is done.

The situation is more complicated with the third sentence, (156c): the sentence is *ambiguous* in that it can be seen as exemplifying either construction, the existential or bodily action, although the interpretation of the sentence may probably be biased, other things being equal, toward the motional sense. From this it can be said that in the sentence (156c), two surface patterns compete with each other, namely [*I walked ... the room*] and [*I ... in the room*], or less specific ones, such as [*I walked ... the ...*] and [*I ... in the ...*] (Table 5.7 and 5.8 show how the two patterns behave in COCA).

We can observe the difference between the prepositions more easily when we see examples with the other prepositions, *into* and *to*:

- (157) a.*I stood into the room.
 b. I danced into the room.
 c. I walked into the room.
 d. I went into the room.

- (158) a.*I stood to the room.

⁷⁾Unfortunately there is no common name for the construction embodied by this type of sentence.

Table 5.7: *X* in [*I walked X ARTICLE*]

rank	X	freq.
1	into	445
2	to	217
3	down	174
4	through	168
5	in	137
6	around	103
7	up	67
8	toward	59
9	along	58
10	across	52

Table 5.8: *X* in [*I X in ARTICLE*]

rank	X	freq.
1	be	7347
2	think	946
3	live	912
4	sit	705
5	believe	533
6	work	427
7	stand	392
8	look	364
9	see	351
10	get	336

b. *I danced to the room.

c. I walked to the room.

d. I went to the room.

In both cases the latter two examples exemplify the intransitive motion construction. This means the examples with the verb *walk* are, unlike (156c), no longer ambiguous. (158a) and (157a) show that if the verb *stand* is used in this context, the unacceptability results. This may be because of the contradiction between the *static* meaning of the verb and the *dynamic* meaning of the preposition *to* and *into*.

Examples with the verb *dance*, (157b) and (158b), are problematic in terms of the following two points: the construction instantiated by (157b) and their difference in acceptability. The construction (157b) instantiates is different from that of (156b): it clearly exemplifies the intransitive motion construction. This suggests the verb *dance* is neutral in terms of the motional sense: it can represent both action and motion, depending on the context in which it is embedded. However, given that (158b) is unacceptable unlike (157b), things are not so simple: if the verb is truly neutral in that respect, the meaning of it should be accommodated, or *coerced*, to the context [*I ... to the room*] and the acceptability should result, but this is not the case, as seen in the examples.

In view of these facts, the following can be concluded: the sequence [*I ... in the room*] cannot be a surface pattern evoking a construction if it appears with some other sequences including verbs with dynamic, motional sense in themselves such as *go*, whereas the similar sequences including *to* and *into* instead of *in* can unambiguously evoke the intransitive motion

construction. This, further, suggests that there can be the difference in the strength of evoking constructions among sequences. Patterns compete with each other and if for a pattern *p* there is another pattern *q* which is stronger than *p* in the same input expression, *p* cannot work as an evoker of any constructions.

In fact this kind of argument has already been done in the previous chapter. What is important here is, however, not the fact that there is such a thing as the difference in the strength of evoking, but *what factor*, or *what part*, is crucial in determining the difference. Evidently in this case, the culprit is the preposition included in a pattern, *in*, *to* and *into*.

There arises a possibility that the preposition *in* itself cannot play a role of surface pattern. This possibility comes from the fact that the constructions instantiated by (156a) and (156b) are different: the difference cannot be reduced to the preposition because the preposition is the same. The crucial part in determining the construction should include the verbs, *stand* and *dance*, lexically or super-lexically. It is probably the case that the super-lexical pattern like [... *danced in* ...] does the job. The pattern of course includes the preposition *in*, but it can only behave as a pattern combined with *danced*; *in* alone can never do.

To verify the possibility, let us see more examples concerning the prepositions. As discussed in 4.4, a prepositional phrase potentially has three possible interpretations in terms of syntactic structure: the caused-motion construction, the simple transitive construction with locative PP and that including a noun phrase composed of a noun and PP. For convenience sentences showed in (83) and (86) are presented again:

- (159) a. I threw the ball in the basket. (= 83a)
 b. I threw the ball in the park. (= 83b)
 c. I saw the ball in the basket. (= 86b)

They differ in their syntactic structures, as represented below by bracketing:

- (160) a. I [threw [the ball] [in the basket]]. (= 84a)
 b. I [[threw [the ball]] [in the park]]. (= 84b)
 c. I [saw [the ball [in the basket]]]. (= 87)

They cannot be differentiated only by the part [*the ... in the ...*].

This fact is clearly related to the evoking strength of the preposition *in*, because thought the sequence [*the ... in the ...*] is indeed a super-lexical one, *in* is only combined with a definite article *the*, which is also seen as a weak evoker. It should be, therefore, beneficial to see how the sequence behaves.

Sequence involving *in*

Now let us consider about the sequence more in deep. There can be three factors affecting the behavior of the sequence, two of which are within it and the other of which is outside of it: the item, in most cases a noun, appearing in the first slot, namely among *the* and *in*; the item, which is also probably a noun, appearing after *in the*; the verb which is assumed to occur just before the sequence.

It is reasonable to assume the following: if the sequence behaves as an evoker of the caused-motion construction, the first slot in the sequence is not so important, because the most important factor to determine the construction it evokes is the noun appearing after *in the* and the verb appearing before that, as seen in the examples just above; if it works as evoker of some noun phrase constructions, in turn, both slots are not important, and hence the verb occurring before it only matters; in consequence, the only difference is whether the second slot matters or not.

This leads to a prediction that if we do not specify what appears in the second slot and only words in the first slot are examined, the sequence cannot behave as a pattern. Hence the pattern we should investigate first is: [ARTICLE X *in*]; in order to obtain larger sample, *the* is generalized into an article, and to include examples including bare nouns and pronouns after the preposition *in*, e.g., *I put the ball in it* and *I put the man in jail*, the second *the* is omitted.

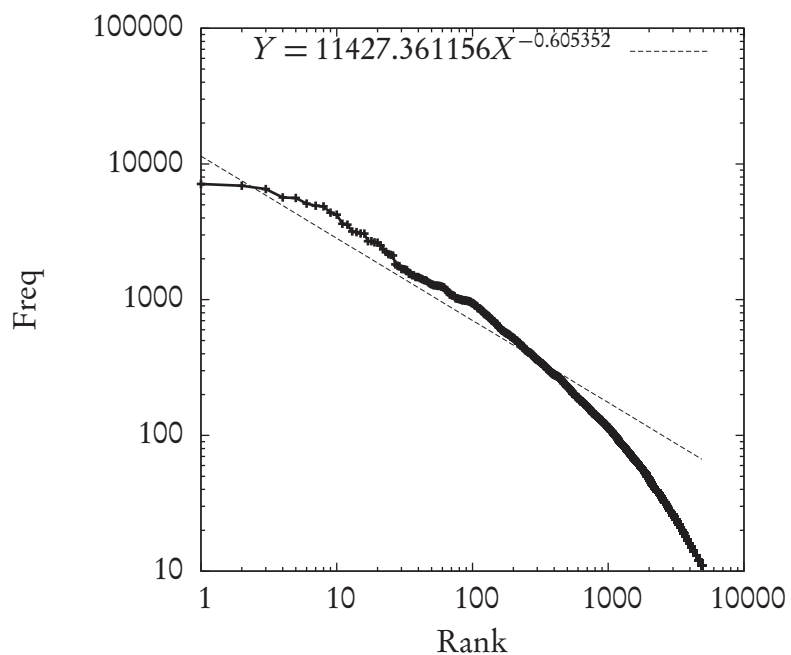
The result is showed in Table 5.9 and 5.12. As predicted, the sequence can hardly seen as functioning a surface pattern. The estimated value of B is far smaller than one (≈ 0.605), and that of A (≈ 11427.36) is far larger than the frequency of the topmost item, *change*, namely 7124.

There are a few things which can be argued based on the data from Table 5.9. The table contains some nouns which can be seen as, in a sense, *requiring* the prepositional phrase headed by *in*, such as *change*, *increase*, *difference* and *interest*. Other nouns are, however, quite general ones in that they are highly frequent and have general or vague meaning: they appear in that slot only because they are popular; they have almost no preference concerning the sequence. This also supports the reasoning presented above.

In order to augment this argument, it is also beneficial to see a pattern with broader context, namely one including the verb position. The pattern is [X PRONOUN *in*]. This pattern can eliminate the possibility of including examples of noun phrase constructions be-

Table 5.9: X in [ARTICLE X in]

rank	X	freq.
1	change	7124
2	man	6912
3	increase	6539
4	way	5665
5	difference	5615
6	people	5110
7	war	4919
8	woman	4867
9	interest	4380
10	year	4233
	\vdots	
	total	580887 ⁺

Figure 5.12: X in [ARTICLE X in]

cause the noun in question is not contained and instead the pronoun is placed at that place. It is assumed that In the X position of the pattern verbs are strongly predicted to appear, and if so, the verbs should be effective factors to determine whether the whole expression is the caused-motion construction or not.

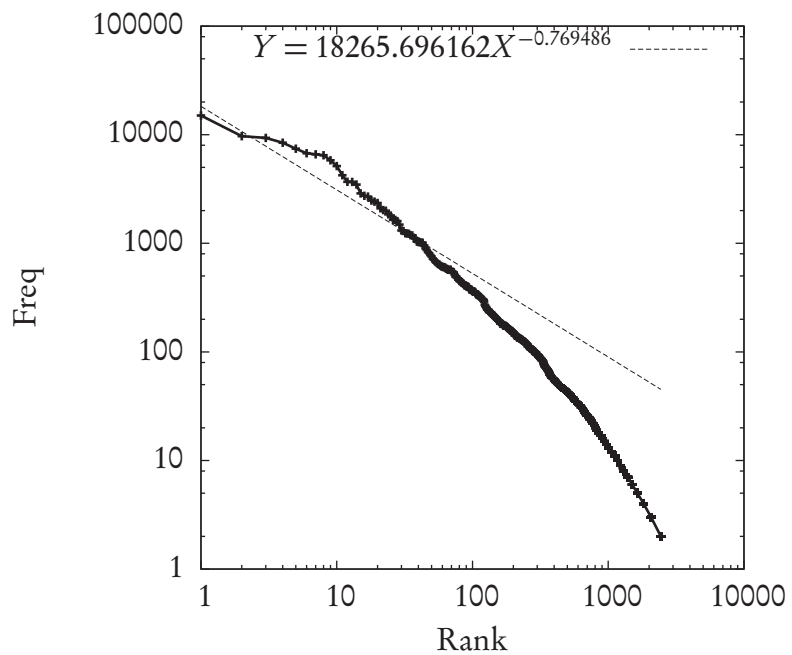
The prediction here is that the words at the position do not show Zipfian distribution and hence the sequence cannot be seen as an evoker of a construction, because the words should include both those mainly used as a part of the caused-motion construction (e.g., *put*) and those not usually used in the construction.

The result is presented in Table 5.10 and Figure 5.10. Contrary to the prediction, the data shows the words in the position are distributed somewhat Zipfianly: the value of A is estimated as around 18625.696, which is close to the frequency of the most frequent word *put*, 15033; the B value is estimated as around 0.77, which is much closer to one compared to the case of [ARTICLE X in] (see Figure 5.12), though it is still a little far from one.

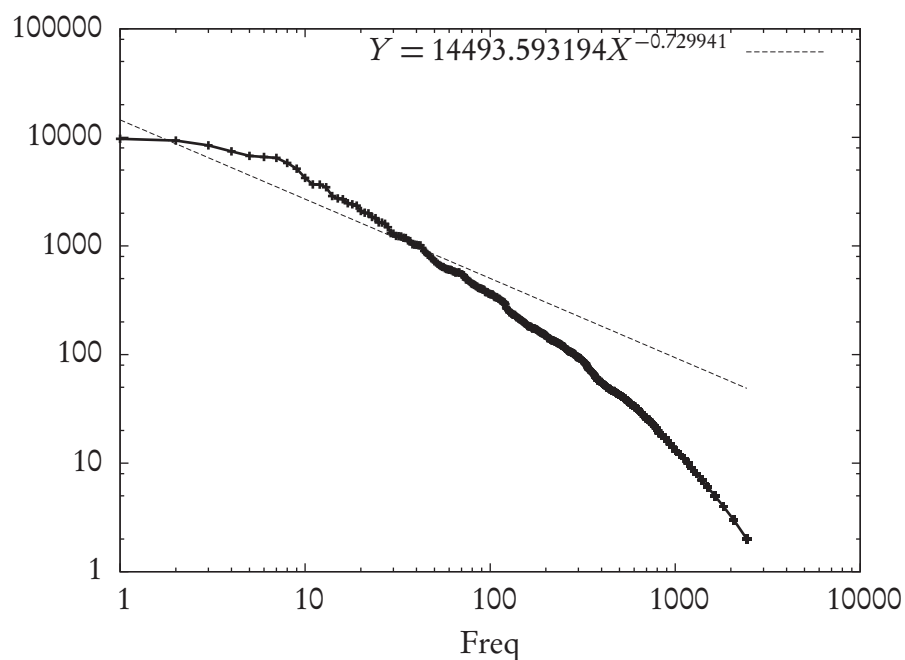
What is remarkable in the result is that the verb *put* occurs extremely frequently in that position, which quite probably make the distribution Zipfian-like. In fact if *put* is eliminated from the list the result falls into a non-Zipfian one (see Figure 5.14): although the B value does not change a lot (≈ 0.73), the A value is estimated as about 14493.59, which is much greater than the frequency of *to*, 9687. This means that the distribution can hardly be seen as obeying

Table 5.10: X in [X PRO-NOUN *in*]

rank	X	freq.
1	put	15033
2	to	9687
3	of	9342
4	see	8445
5	be	7437
6	find	6748
7	for	6604
8	with	6478
9	do	5820
10	,	5155
	⋮	
total		245773+

Figure 5.13: X in [X PRONOUN *in*]

power law.

Figure 5.14: X in [X PRONOUN *in*], where *put* $\notin X$

Back to the original result including the verb *put*, it seems possible to conclude that the sequence can behave as an evoker of the caused-motion construction, because it strongly associates the exemplars including the verb *put* appearing at that position. This is, however,

probably not the case. As you can see there are a number of prepositions in the top ten words, *to*, *of*, *for* and *with*, and the total counts of those prepositions, 32111, is well over the frequency of *put*, 15033. This means the exemplars associated with the sequence are not consistent in their meaning, which makes it far less effective on construction evoking (see 3.3.9 about the *non-contradictory condition*).

Form these it can be concluded that the sequence cannot evoke any constructions in most cases, but, if there are no competing sequences evoking any other constructions, it may work as an evoker of the caused-motion construction. The final point is also a prediction about the behavior of it and hence testable with some experimental approach, which is one of the future tasks.

Sequence involving *into* and *to*

Let us move on to the behavior of the other two prepositions, *into* and *to*. As seen in the preliminary part of this subsection, the two prepositions are considered to have relatively strong preferences to some specific constructions, compared to *in*. From this it follows that we should investigate the behavior of the similar sequences to those examined above, in this case including *into* and *to* instead of *in*, namely [ARTICLE X *into*], [X PRONOUN *into*], [ARTICLE X *to*] and [X PRONOUN *to*].

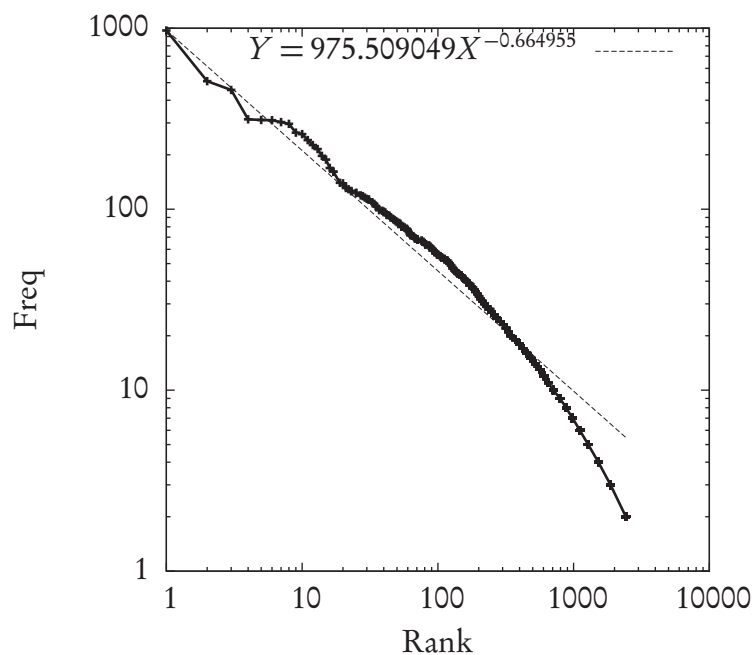
The prediction about their behavior is the following: as for the first and the third sequence, items at the slot do not show Zipfian distribution, while for the second and the forth, items at that position are distributed almost Zipfianly. The first and the third sequence indeed have an ambiguity about whether they are related to the caused-motion construction or some noun phrase constructions, but the second and forth ones do not, because the existence of pronouns at the place just before the prepositions eliminates the possibility of being a part of any noun phrase constructions.

As for *into*, the results are shown in Table 5.11 and Figure 5.15 as to the sequence [ARTICLE X *into*], and in Table 5.12 and Figure 5.16 as to the sequence [X PRONOUN *into*]. In a word, the prediction is confirmed: in the former sequence, despite the estimated *A* value (≈ 975.51) being really close to the frequency of the topmost words, 972, the *B* value is estimated as small as about 0.66; in the latter, the *A* value (≈ 4474.45) is near the frequency of topmost words, *turn* (3886), and the *B* value is also close to one (≈ 0.81).

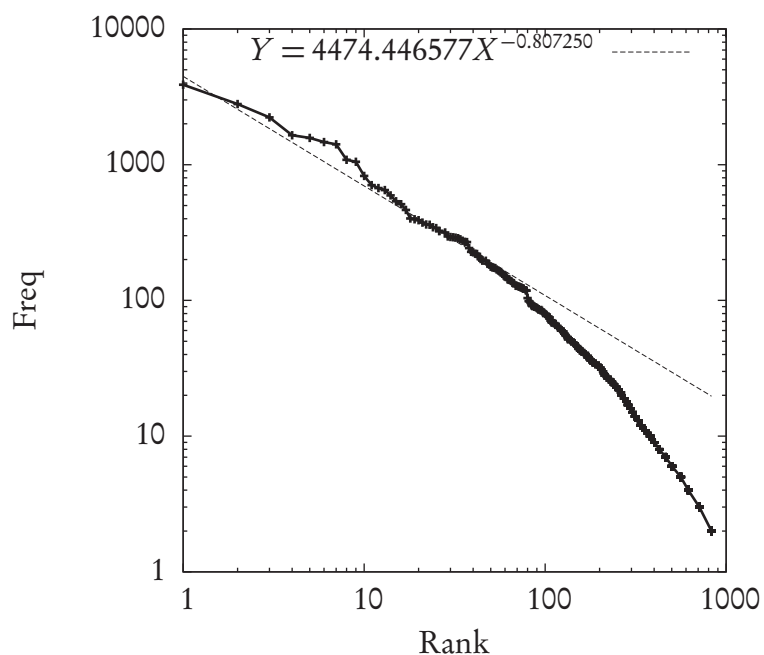
However, the sequences including *to* do not behave as expected (see Table 5.13 and Figure 5.17). As seen in Figure 5.17, the estimated *B* value is unexpectedly high (≈ 0.74); *A* value is

Table 5.11: X in [ARTICLE X into]

rank	X	freq.
1	investigation	972
2	way	510
3	window	456
4	country	314
5	glimpse	311
6	border	310
7	world	303
8	car	296
9	inquiry	265
10	ball	260
	\vdots	
total		38128

Figure 5.15: X in [ARTICLE X into]Table 5.12: X in [X PRONOUN into]

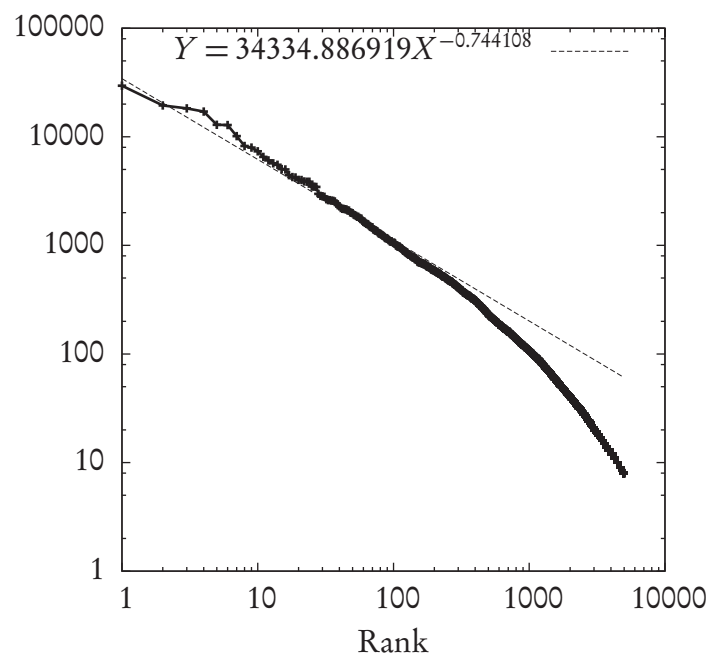
rank	X	freq.
1	turn	3886
2	get	2796
3	put	2233
4	make	1646
5	take	1574
6	bring	1468
7	throw	1414
8	transform	1087
9	lead	1051
10	talk	826
	\vdots	
total		

Figure 5.16: X in [X PRONOUN into]

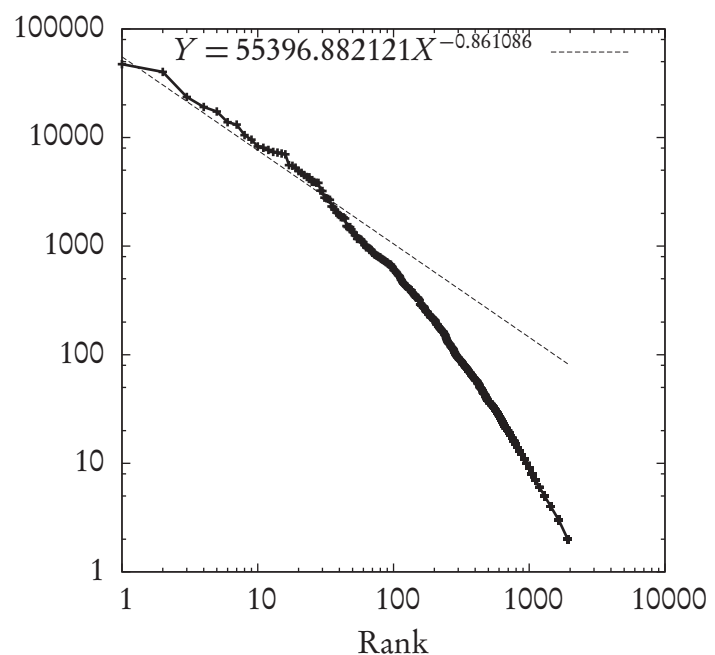
estimated to be not so far from the frequency of the topmost item, 29857. This means that the sequence [ARTICLE X to] can be regarded as an evoker of some construction, in addition to [X PRONOUN to], which can no doubt be seen as a surface pattern due to the estimated values of A and B (see Table 5.14 and Figure 5.18).

Table 5.13: X in [ARTICLE X to]

rank	X	freq.
1	way	29587
2	chance	19453
3	right	18288
4	opportunity	17089
5	need	12924
6	ability	12818
7	effort	10158
8	time	8234
9	attempt	7958
10	key	7349
	\vdots	
total		732828 ⁺

Figure 5.17: X in [ARTICLE X to]Table 5.14: X in [X PRONOUN to]

rank	X	freq.
1	for	47383
2	want	40207
3	allow	23620
4	have	19135
5	ask	17360
6	tell	13869
7	take	13154
8	be	10526
9	get	9553
10	bring	8285
	\vdots	
total		

Figure 5.18: X in [X PRONOUN to]

This is most probably because of another function of to , that is, to mark a *to-infinitive*: presumably the sequences strongly evoke the constructions involving *to-infinitive*, hence the good fitness to Zipf's law. The difference between the two sequences can be explained in such a way that, for the former sequence, [ARTICLE X to], though noun phrase constructions

including to-infinitive as in, for example *I got a chance to see her*, is predominant, there remain two other possibilities of evoking, namely the caused-motion construction and what can be called the causative to-infinitive construction such as *I asked the student to come*; for the latter sequence, [X PRONOUN *to*], since the possibility of evoking noun phrase constructions are eliminated, there are only two possibilities, resulting in the better fitness.

In fact, taking a look at the result, all the words in Table 5.13 are found to have strong tendency to precedes to-infinitive, or put otherwise, they are such words as require to-infinitive to follow them. Taking the second most one, *chance*, as an example, if we examine words appearing after the sequence [ARTICLE *chance to*] on COCA, the list of words is filled with hundreds of verbs (e.g., *see, get, be, talk, do, win, make, go, play, work* and so on). Similarly, words in Table 5.14 include many verbs which is preferredly used in the causative to-infinitive construction such as *I asked her to come with me*. The most notable about the list is that the topmost item is not a verb, but the preposition *for*, which is, in this context, probably used as a marker of notional subject as in *for you to come*.

To eliminate the possibility of evoking to-infinitive related constructions, it should be modified to include an article after the preposition, namely [ARTICLE X *to* ARTICLE] and [X PRONOUN *to* ARTICLE]. If *to* is followed by an article, it is impossible for *to* function as to-infinitive marker; it should be a preposition.

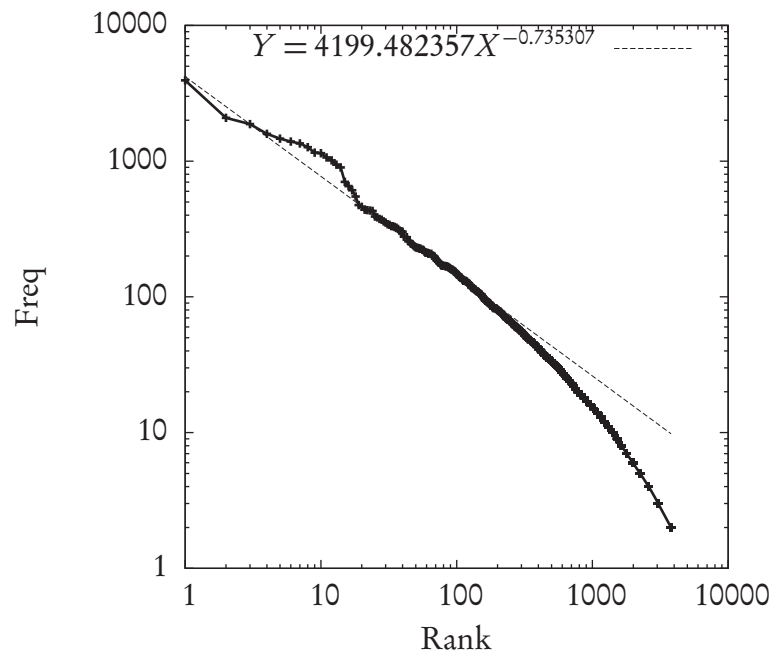
The results obtained with the modified sequences are shown in Table 5.15 and Figure 5.19 for [ARTICLE X *to* ARTICLE], and in Table 5.16 and Figure 5.20 for [X PRONOUN *to* ARTICLE]. In the case of the latter sequence, the estimated values of parameters *A* and *B* show really good fitness: *A* is almost equal to the frequency of the topmost item and *B* is close to one (see Figure 5.20), which is as expected. However, for the former sequence, the result cannot be said as expected: the degree of fitness is almost the same as that of [ARTICLE X *to*]; that is to say, the difference between *A* value and the frequency of the topmost word and the closeness of *B* value to one are not so different from those of [ARTICLE X *to*].

This is probably because, seeing the words listed in Table 5.19, they can be seen as such words as requiring to-phrase after them to specify the meaning of the words. In other words, the preposition *to* is considered to be *noun-oriented*, as opposed to *verb-oriented*; the noun-orientedness of *to* (phrase) may probably decrease the degree of structural ambiguity of the sequence, resulting in the relatively good fitness.

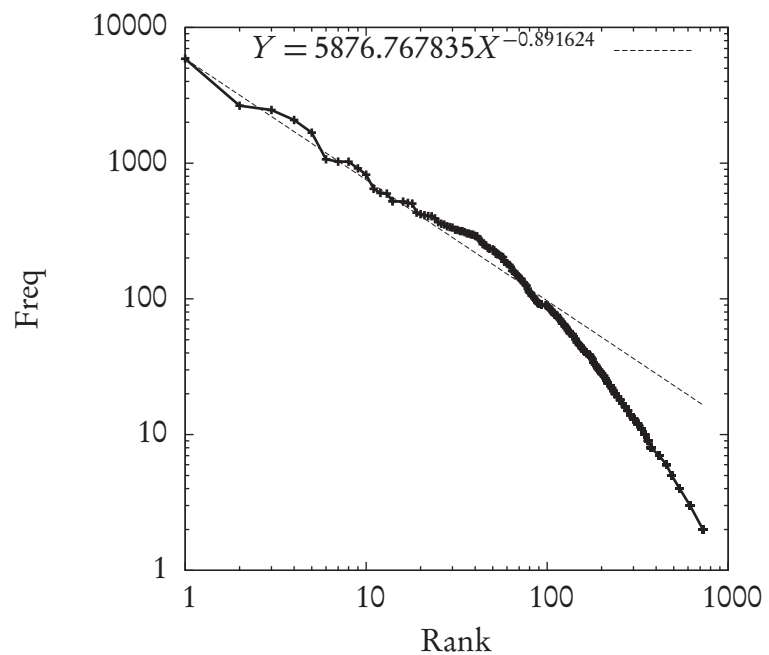
This reasoning can be supported to some degree if we examine what kind of words pre-

Table 5.15: X in
[ARTICLE X to
ARTICLE]

rank	X	freq.
1	way	3940
2	door	2086
3	key	1876
4	end	1588
5	letter	1464
6	answer	1394
7	trip	1350
8	entrance	1267
9	threat	1155
10	visit	1145
\vdots		
total		

Figure 5.19: X in [ARTICLE X to ARTICLE]Table 5.16: X in [X PRO-
NOUN to ARTICLE]

rank	X	freq.
1	take	5881
2	bring	2652
3	make	2466
4	lead	2080
5	send	1678
6	introduce	1068
7	get	1027
8	give	1024
9	with	915
10	drive	822
\vdots		
total		52037

Figure 5.20: X in [X PRONOUN to ARTICLE]

cede the sequence. Table 5.17 shows the list of words occurring before the sequence, where $X' = \{\textit{way}, \textit{door}, \textit{key}, \textit{end}, \textit{letter}, \textit{answer}, \textit{trip}, \textit{entrance}, \textit{threat}, \textit{visit}\}$, that is, the top ten words appearing in the X slot of [ARTICLE X to ARTICLE]. Items appearing in the list are quite

different from those in Table 5.16. There are only a few verbs in it and the few verbs, *be*, *open* and *put*, are not included in Table 5.16.

Table 5.17: *Y* in [Y ARTICLE X' to ARTICLE]

rank	X	freq.
1	all	2080
2	on	1369
3	be	1170
4	open	708
5	in	532
6	.	525
7	at	517
8	,	492
9	put	446
10	for	423

Problems with pronouns

It seems that the results presented above show the effectiveness of those sequences with the preposition *into* and *to* in the context of [PRONOUN ... (ARTICLE)]. However, taking a close look at the data, we can easily find that the results include several different constructions with the same sequence. In other words, the sequences are still *ambiguous* to some degree. For example, the list in Table 5.16 includes the verb *make*, which is hardly seen as a part of the caused-motion construction. In fact, examples embodying the pattern [*make* PRONOUN to ARTICLE] almost necessarily contain *it* as the PRONOUN position, such as *He made it to the top*. This is, clearly, not an instance of the caused-motion construction.

Fortunately there is a possibility that the still remaining ambiguity can be somewhat eliminated by specifying pronouns at the PREPOSITION position. As for the above raised example of [*make* PRONOUN to ARTICLE], it can be eliminated if the pronoun is specified with other than *it*. Generally speaking, the preposition *it* behaves rather differently from other prepositions because of *its* (non-)animacy and the expletive usage.

Animacy is known to have effects on grammatical behavior of an expression or construction. For example, *take* and *give* have different preferences as to their pronoun object in the context of [*take/give* PRONOUN to ...]. Both verbs can be used with an *inanimate* pronoun object without any awkwardness:

(161) a. He took it to school.

b. He gave it to me.

However, the latter, *give*, sounds somewhat awkward if used with animate pronoun object in the context:

(162) a. He took her to the station.

b. ?He gave her to me.

This does not mean that (162b) is unacceptable, but is somewhat bizarre, which probably needs some context by which the expression becomes appropriate.⁸⁾

Consequently, the queries should be modified with specified pronoun(s) and searched again. For this purpose pronouns are divided into four categories: $PRO_1 = \{me, you, him, her\}$; $PRO_2 = us$; $PRO_3 = it$; $PRO_0 =$ the others. The category PRO_1 is that of singular animate pronouns, though *you* is ambiguous about the number. The category PRO_3 has only one member, *it*, which is intended to represent the behavior of inanimate pronouns. The category PRO_2 is also a one-member set, which is offered in order to see whether the behaviors of words at the slot of [X PRONOUN *to*] according to the number of pronoun.

PRO_0 includes *them*, reflexive pronouns and indefinite pronouns. They are, however, not examined because *them* is ambiguous about animacy, reflexives can be used adverbially, indefinites can be used as parts of noun phrase construction such as *something to the effect . . .*

The results are provided in Table 5.18-5.20 and Figure 5.21. Generally, for all the three sequences they show really good fitness to Zipf's law: all the *B* values are estimated to be almost equal to one, and *A* values are also close to the frequencies of the topmost items, though the *A* value in the PRO_2 condition (≈ 658.495) is relatively far from the frequency of *bring* ($= 559$).

There are a few findings comparing the results among the three conditions. As expected, the topmost word in the PRO_3 condition is *make*, which, in turn, does not even appear in the top ten lists of the other two conditions. As for the effect of number, comparison between the PRO_1 and the PRO_2 conditions makes almost no remarkable difference: seven out of ten items are shared with the two lists and the member of top three items is the same, despite the difference in its order.

⁸⁾In fact, in COCA we can find 255 examples of [*give me/you/her/him to*] where *to* is specified as a preposition. Incidentally, the number of hits of [*give it to*] is 4973.

Table 5.18: X in
[X PRO₁ to ARTICLE]

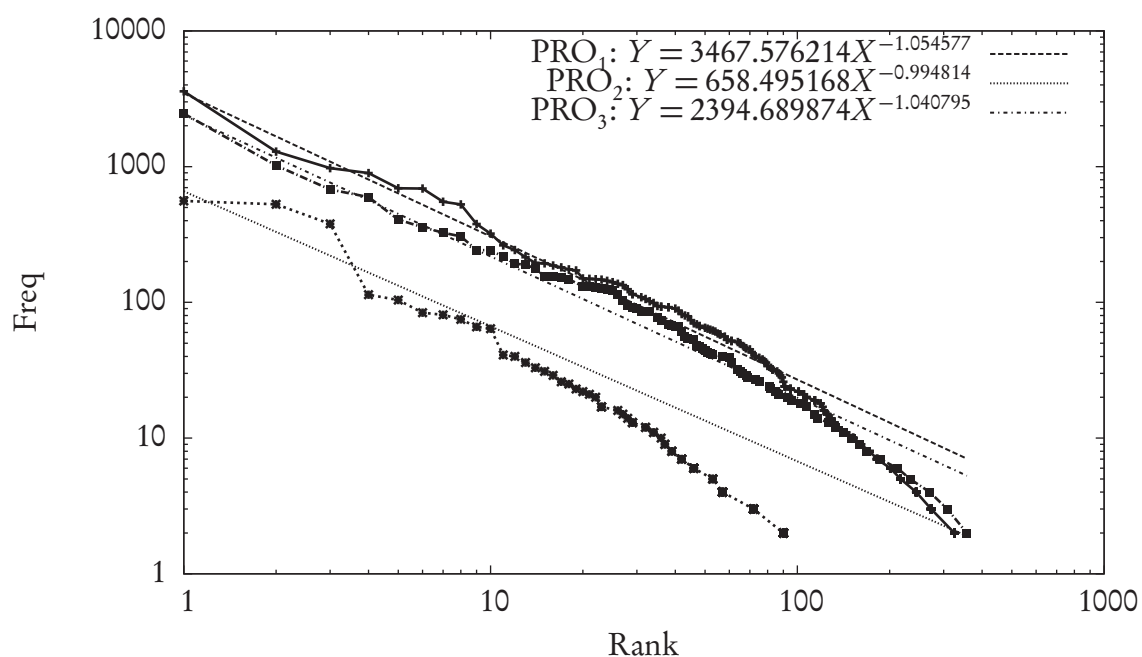
rank	X	freq.
1	take	3590
2	lead	1292
3	bring	975
4	send	896
5	with	692
6	introduce	690
7	drive	551
8	get	527
9	follow	378
10	invite	321
	⋮	
total		19244

Table 5.19: X in
[X PRO₂ to ARTICLE]

rank	X	freq.
1	bring	559
2	take	528
3	lead	379
4	introduce	114
5	get	104
6	return	84
7	drive	81
8	alert	75
9	of	66
10	with	64
	⋮	
total		3012

Table 5.20: X in
[X PRO₃ to ARTICLE]

rank	X	freq.
1	make	2461
2	take	1021
3	give	679
4	bring	597
5	send	407
6	compare	358
7	leave	326
8	sell	306
9	hand	243
10	return	242
	⋮	
total		14580

Figure 5.21: X in [X PRO _{i} to ARTICLE]

Some words are shared with all the conditions, namely *take* and *bring*, and other words in PRO₃ condition are also contained either in PRO₁ or PRO₂ conditions, that is, *send* and *return*,⁹⁾ and three of them, *take*, *bring* and *send*, are included in the top five of the words appearing X slot of [X PRONOUN to ARTICLE] presented Table 5.16. This means that

⁹⁾*send* is ranked 15th in the PRO₂ condition and *return* is ranked 29th in the PRO₁ condition.

those words can be used regardless of the type of pronoun, be it animate or inanimate, and singular or plural. In fact, The three words also appear within the top-five in the words at the *X* slot if the pronoun is specified with *them* (see Table 5.21).

Table 5.21: *X* in [*X them to* ARTICLE]

rank	X	freq.
1	take	659
2	bring	378
3	lead	370
4	send	287
5	get	166
6	of	158
7	return	148
8	give	142
9	add	130
10	carry	116

In this connection, it can be assumed that the examples obtained from the research still include two different types of construction, one of which is, of course, the caused-motion construction. The other is what Goldberg (1995:90) calls the *transfer-caused-motion construction*, as exemplified by the following:

(163) a. The judge awarded custody to Bill

b. Bill gave his house to the Moonies.

(Goldberg 1995:89)

This construction is explained as having the meaning of transfer of ownership (Goldberg 1995:89), not implying a physical motion, at least primarily. The typical example of the construction is one including the verb *give* as in (163b), which is often argued in the context of what is called the *dative alternation*, an alternation between the ditransitive construction and the transfer-caused-motion construction:

(164) Bill gave his house to the Moonies. ⇔ Bill gave the Moonies his house.

(Coleman & De Clerck 2009:16)

While Goldberg (1995:89-90) argues that the transfer-caused-motion construction is a kind of metaphorical extension from the simple or physical caused-motion construction and perhaps regards it as a subtype of the caused-motion construction, some claims that those two

are qualitatively different and the latter is a kind of *caused-possession* construction, not of the caused-motion construction (e.g., Coleman & De Clerck 2009). In the framework of the current theory, EBCG, they should be distinguished by means of surface patterns. If the second position is taken, which is probably possible.

The key to the distinction is: the verb. It is reasonable to assume that the verbs which can be used in the transferred-caused-motion construction is rather limited, at least compared to those usable in the caused-motion construction, so it is predicted that the sequence such as [... *it to* ...] can be seen as an evoker of the caused-motion construction, but if the verb is specified as in [*give it to* ...], the transferred-caused-motion construction is evoked.

5.2.4 Patterns on major English constructions

Since we have already grasped the general tendency of the distributional behavior of a pattern, this is the time to investigate patterns involving specific constructions. The constructions examined here are those analyzed in the previous chapter, namely the *ditransitive construction*, the *resultative construction*, the *caused-motion construction*, and the *way construction*.

The ditransitive construction

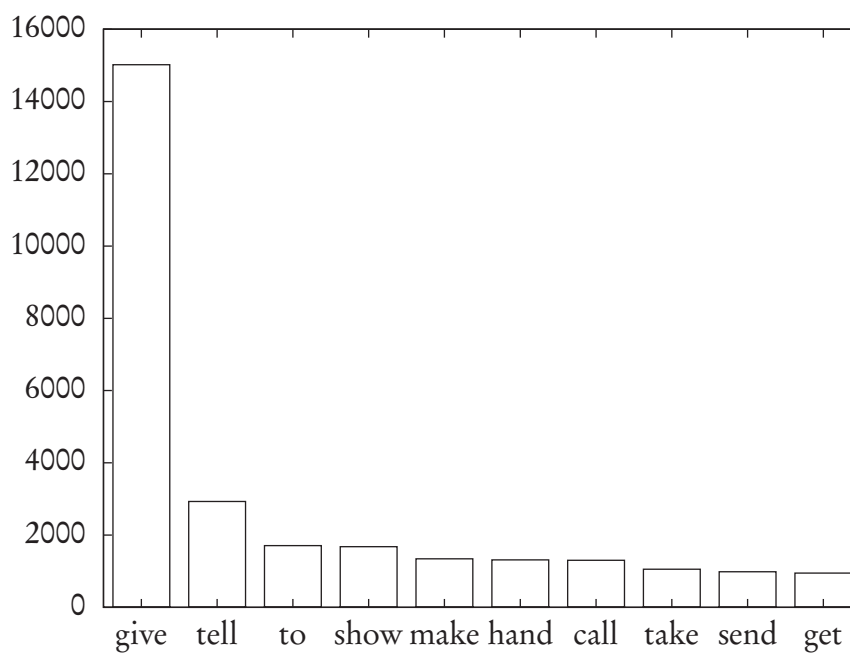
As a first example, a pattern related to the ditransitive construction is taken. As discussed in the previous chapter (4.2.1), a short and simple sequence [PRONOUN ARTICLE] can be a surface pattern evoking the construction. Here, by checking whether the pattern has a Zipfian nature of frequency distribution in words appearing before and after the pattern, the validity to assume it as a plausible surface pattern to evoke the construction.

For this purpose, words appearing just before the sequence are searched. The pattern examined here is, therefore, [X PRONOUN ARTICLE]. However, for some technical limitations (see fn. 2)) on the corpus interface of COCA, the exact searched sequence specifies the pronoun as *me*, hence [X *me* ARTICLE]. The ARTICLE positions are specified by part of speech tags. The result is provided in Table 5.22, Figure 5.22 and Figure 5.23. The first two show the words which are ranked in the top ten; the third one is a log scaled graph containing words occurring more than once.

From Table 5.22 it is clear that a verb *give* occurs extremely frequently in that position. Its frequency is more than one third of the total frequency, which amounts to nearly 37%. The overall tendency of frequency distribution can be seen as Zipfian, as seen in the Figure

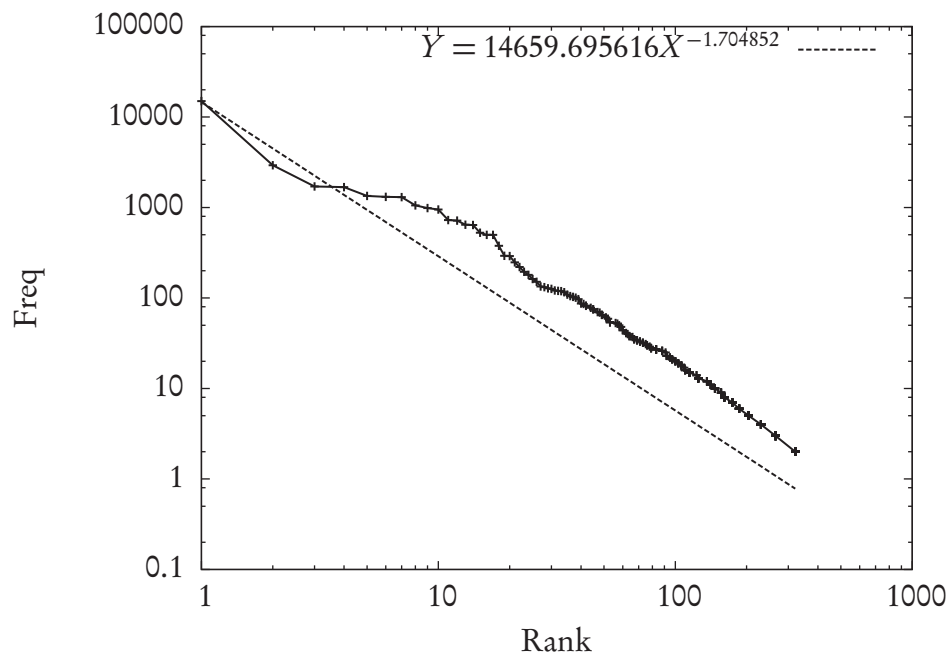
Table 5.22: X in $[X \text{ } me \text{ ARTICLE}]$

rank	X	freq.
1	give	15017
2	tell	2929
3	to	1710
4	show	1679
5	make	1344
6	hand	1313
7	call	1303
8	take	1056
9	send	986
10	get	950
	\vdots	
total		40839

Figure 5.22: Histogram of X in $[X \text{ } me \text{ ARTICLE}]$

5.23. The value of B is estimated as far greater than 1, namely $1.704852\dots$, and the value of A , $14659.695616\dots$, is really close to the frequency of the topmost word *give*, 15017.

It may be a problem that we can find a word which does not in any sense constitute the ditransitive construction if appearing in that position, namely a preposition *to*. Actually there are eight prepositions in the top 100 words (see Table 5.23). This is not a problem, however,

Figure 5.23: X in $[X \text{ me ARTICLE}]$

because we can assume that patterns such as $[to \text{ PRONOUN ARTICLE}]$ are evokers of some other construction(s).

Table 5.23: Prepositions appearing in X

rank	preposition	freq.
14	for	641
18	with	377
23	at	195
26	in	151
37	on	103
42	of	82
53	from	54
100	about	20

The resultative construction

The second example of the construction examined is the resultative construction. As seen in 4.3.1, for the resultative construction it is almost impossible to identify a general surface pattern to specify the examples of the construction, and the construction is best characterized by a *mosaic* of semi-fixed phrases including verb-predicate pairs such as *shoot ... dead* and *tear*

...*apart* and noun-predicate pairs such as *door open*. Almost the same thing, though their conclusion is somewhat different, is claimed by Goldberg & Jackendoff (2004) (see 4.3.3).

In consequence, here two patterns related to the construction are investigated as a set of case studies. They are:

- (165) a. [... DETERMINER NOUN to death]
 b. [... DETERMINER head off]

The former sequence includes an prepositional phrase usable as an resultative predicate (RP) of the resultative construction, *to death*, which is a member of the RP list presented by Boas (2003) as seen in (61). The prepositional phrase itself is considered to behave as a surface pattern of the resultative construction (see 4.3.2), and it is probably beneficial to examine how items before the sequence distribute. The exact sequence examined here is, however, not the prepositional phrase itself, but [... DETERMINER NOUN to death], namely the prepositional phrase with preceding [DETERMINER NOUN]. The sequence is added because the original prepositional phrase can function either as the resultative phrase of *intransitive* or *transitive* sentence, as exemplified below, which may probably make very noisy the frequency distribution of words appearing just before the sequence:

- (166) a. He froze to death.
 b. She stabbed the man to death.

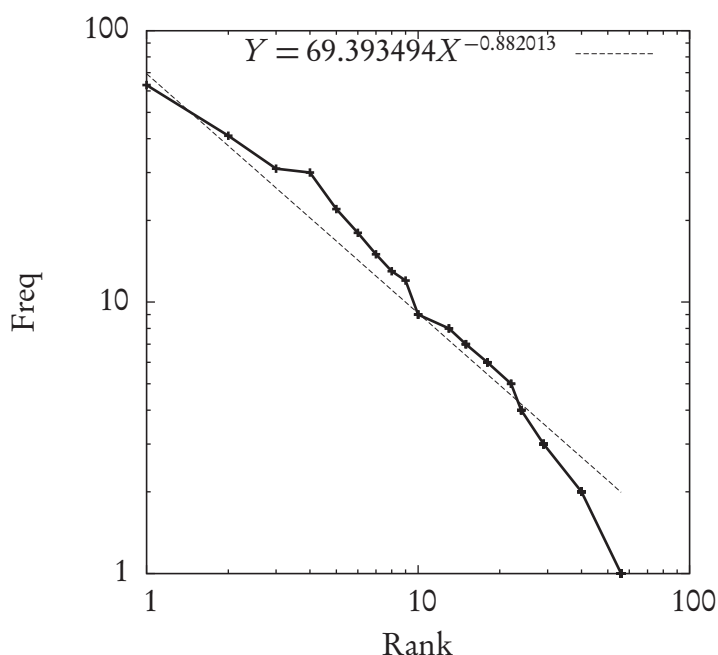
In cases where [DETERMINER NOUN] precedes the prepositional phrase, the whole sequence is highly likely to be transitive, as seen in (166b). The category *NOUN* is used here despite being an open-class category for convenience sake.

The result is as shown in Table 5.24 and Figure 5.24. The obtained distribution can be interpreted as Zipfian: the *A* value is almost the same as the frequency of the topmost item, *beat*, and the *B* value is near one. There are a few problematic items in Table 5.24, namely *of* and *and*, influences of them on the distribution recognition is considered to be subtle because of their frequencies. Other frequent words can be seen as what typically consist of resultative sentences as main verbs, such as *beat* and *stab*, as seen in 4.3.1.

The latter sequence, [X DETERMINER head off], includes a sequence found to be frequently contained in resultative clauses, namely [... head off] (see 4.3.1, especially Table 4.8), as exemplified below (underline added):

Table 5.24: X in [X DETERMINER NOUN to death]

rank	X	freq.
1	beat	63
2	stab	41
3	love	31
4	shoot	30
5	put	22
6	of	18
7	scare	15
8	choke	13
9	and	12
10	bludgeon	9
10	send	9
10	sentence	9
⋮		
total		527

Figure 5.24: X in [X DETERMINER NOUN to death]

- (167) a. Then, taking a deep breath, began to bawl its head off. (=65e)
 b. We all chat our heads off don't we? (=65f)

Head in many cases follows some determiner, so X in the sequence [X DETERMINER head off] is considered to represent the position at which a main verb of a resultative clause appears and the frequency distribution of items at that position is expected to obey Zipf's law.

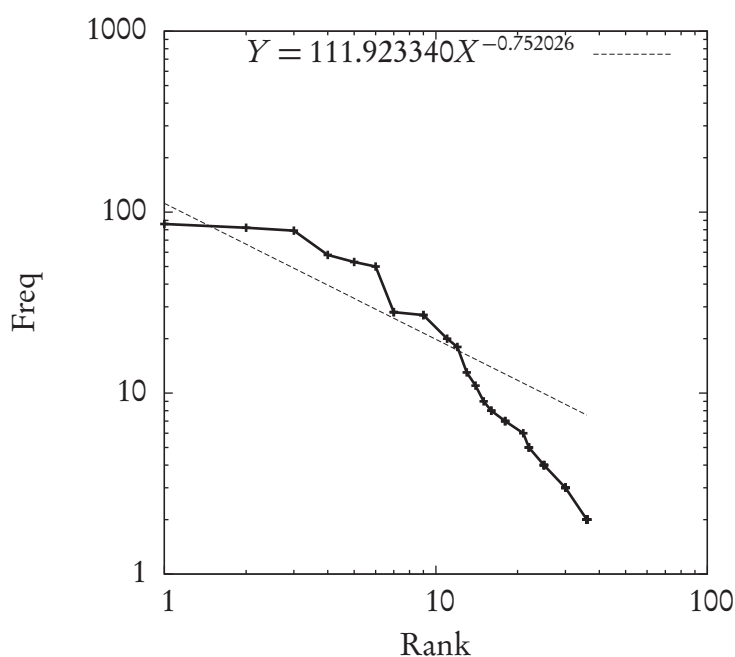
The result is, however, not as expected. Table 5.25 and Figure 5.25 show that both of the A and B values present deviations from Zipfian distribution. The top three items, *blow*, *bite* and *cut*, are almost the same in terms of their frequency, and hence the graph is far from linear.

There is, however, a possibility to interpret this result as a counter-evidence for the sequence [... head off] functioning as a surface pattern. Looking again at the examples including [... head off] such as (167a=65e) and (167b=65f), we find a large number of examples in which a *possessive determiner* (e.g., *its*, *our*) precedes the noun *head*. In fact out of 104 examples with [... head off] sequence in Boas's (2003) data retrieved from BNC, as many as 94 are used with possessive determiners. This means that if the *DETERMINER* position is specified as a possessive determiner, the distribution may change and become Zipfian.

This is partly the case, as seen in Table 5.26 and Figure 5.26. Although the A value is a

Table 5.25: X in [X DETERMINER head off]

rank	X	freq.
1	blow	86
2	bite	82
3	cut	79
4	scream	58
5	laugh	53
6	take	50
7	tear	28
7	chop	28
9	lift	27
9	rip	27
	\vdots	
total		759

Figure 5.25: X in [X DETERMINER head off]

little larger than the frequency of the topmost word, the difference is much smaller than that of the case of [X DETERMINER head off], and B value also becomes closer to one. The still remaining deviations can be due to the mixture of the functions of *off* just after *head*, in the sense that *off* at that position can function both as a particle and as a preposition, the latter of which takes another argument as its object and may form another construction, namely the *caused-motion* construction. In fact all the few examples with non-possessive determiner in Boas (2003) data instantiate the caused-motion construction (indices are the same as those presented in the previous chapter; underline added):

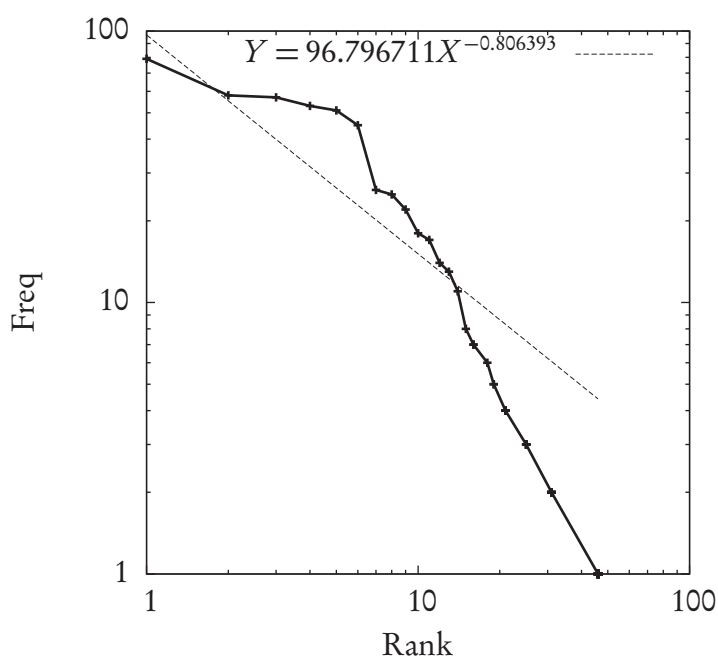
- (168) a. The wretched animal could only carry away one chicken but he
bit the heads off all the others in a sheer lust of killing. (BOAS:off-178)
- b. This Katherine bites the heads off rag-dolls and threatens her sister Bianca with a pair of pinking shears. (BOAS:off-228)

This makes a sharp contrast with the sentences presented in (167a=65e) and (167b=65f).

Clearly, expressions with the sequence whose X slot is filled with some of the frequent words found in Table 5.26 such as *blow POSS head off* and *scream POSS head off* (but not *cut POSS head off*) have somewhat idiomatic senses which are hard to expect from the composing words. This means that sequences including those phrases may become ambiguous,

Table 5.26: X in [X POSS head off]

rank	X	freq.
1	blow	79
2	scream	58
3	cut	57
4	laugh	53
5	bite	51
6	take	45
7	lift	26
8	tear	25
9	rip	22
10	chop	18
	⋮	
total		622

Figure 5.26: X in [X POSS head off]

and therefore that the sequence [X POSS head off] itself cannot be regarded as a surface pattern, at least in terms of the frequency distribution of the word appearing at the X position.

The caused-motion construction

As for the caused-motion construction, we have already seen the behaviors of some sequences including major prepositions in terms of the distributional features of words at a specific position in the sequences (5.2.3). Therefore here a sequence analyzed in 4.4.2 is investigated, which is included in a single sentence which seems difficult to deal with under the exemplar-based framework, namely the following:

(169) Joe squeezed the ball inside the jar. [= (98)]

The sequence investigated is [X ARTICLE ball inside], which is considered to play an important role as an evoker of the caused-motion construction. Examples with the sequence are as follows (from COCA; underline added):

- (170) a. Normally you want to get the ball inside and get fouled, ... (COCA:NYTimes) [= (103a)]
 b. I was throwing the ball inside. (COCA:Atlanta) [= (103c)]

The frequency table of the words appearing just before *ARTICLE ball*, represented as *X* in [X ARTICLE ball inside], is already provided in Table 4.16 in 4.4.2, but, for convenience, is provided here again in a somewhat modified form (Table 5.27), with a plot graph of rank-freq distribution (Figure 5.27). As is clear from the table and graph, the frequency distribution can be judged as Zipfian. As discussed in 4.4.2, in the table there is a word which cannot be regarded as composing a part of caused-motion clause, that is, *have*, but since its freq frequency is very low both relatively and absolutely, its influence on the whole distribution is considered to be scarce.

Table 5.27: *X* in [X ARTICLE ball inside]

rank	X	freq.	Freq
1	get	16	
2	pound	9	
3	have	3	
4	pump	2	
4	run	2	
4	dribble	2	
4	work	2	
⋮			
total		52	

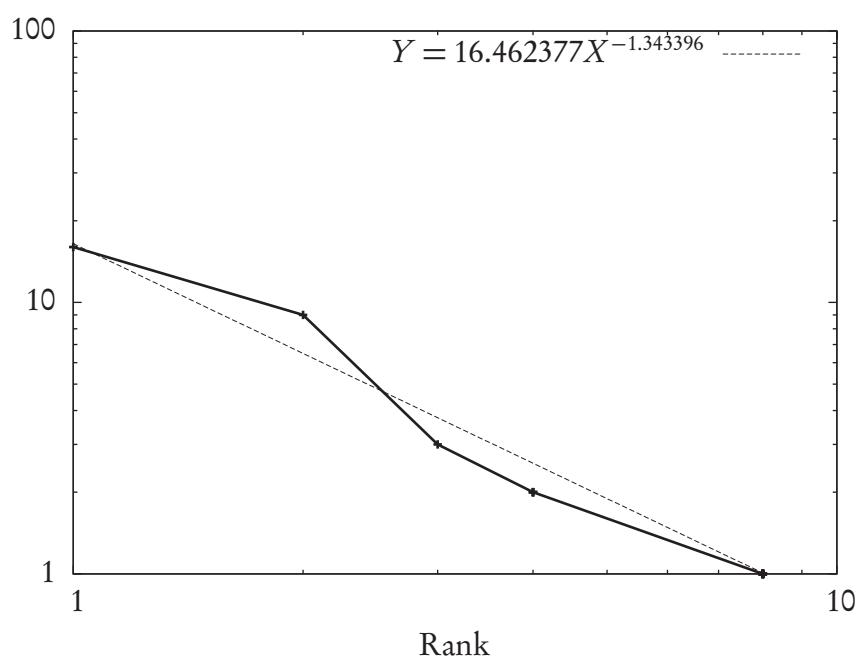


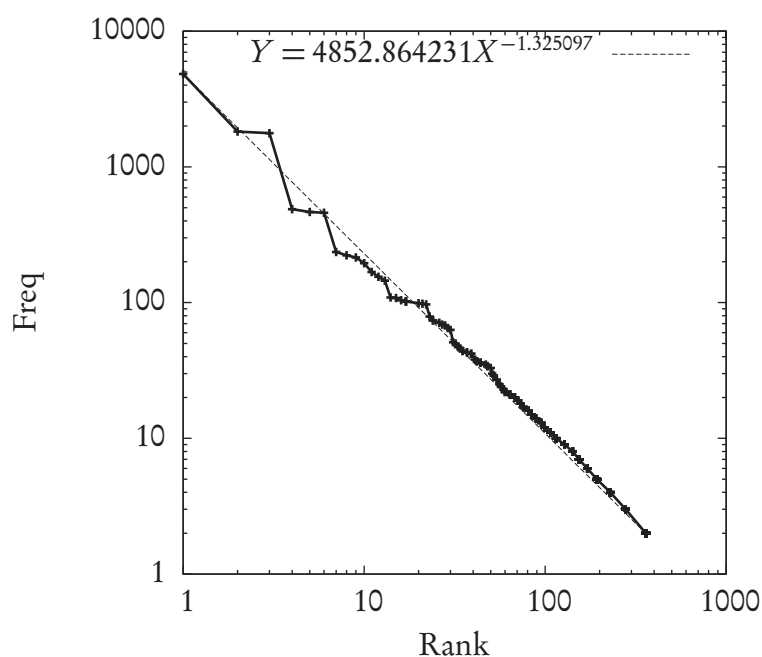
Figure 5.27: *X* in [X ARTICLE ball inside]

The way construction

As seen in 4.5.1, the English *way* construction can be specified by a pattern [... POSS way PREPOSITION], although there are explainable exceptions such as a pattern in which prepositions precede the sequence (i.e., [PREPOSITION POSS *way* PREPOSITION] such as *on my way to*). Therefore it is expected that words at *X* position of [X POSS way PREPOSITION] are considered in most cases verbs and their frequency distribution obeys Zipf's law, which is actually the case: Table 5.28 and Figure 5.28 show that the distribution is Zipfian, with the *B* value of the regression equation much larger than one.

Table 5.28: X in [X POSS *way* PREPOSITION]

rank	X	freq.
1	make	4832
2	work	1819
3	find	1773
4	push	488
5	pick	465
6	fight	459
7	talk	236
8	feel	223
9	force	215
10	wind	195
	⋮	
total		16638

Figure 5.28: X in [X POSS *way* PREPOSITION]

5.3 Analysis of child-directed speeches

Large-scaled balanced corpora are actually suitable to see some general tendency of linguistic phenomena because of their relatively less sparseness, which leads to the representativeness of input data we humans are assumed to obtain. When it comes to language learning, however, some problems arise as to a couple of characteristics they have: the *genres* or *styles* contained by them and the largeness of their scale.

For most of the large-scaled corpora today in general, and the corpus employed in the current study, COCA, in particular, the major component of each corpus is the written texts, as opposed to spoken. As shown in Table 5.3, the spoken data only amounts to about 20% of all the data contained in COCA ($= 95,385,672/464,020,256$). This imbalance is due to the assumption that the spoken language is only one variation of human language and, therefore, should be treated in the same way as the other types of texts such as academic texts and novels. It cannot be said that the assumption itself is wrong; in order to investigate some general tendencies of a language, that kind of evenness of distribution is in many cases necessary.

However, this is not probably the case when we investigate issues on language learning, especially by children. In most cases the type of language children are exposed to is not the written one, but the spoken one, whose characteristics should be drastically different

from those of, for example, novels and academic texts. Yet the two genres, FICTION and ACADEMIC, constitute almost 40% of COCA, amounting to twice as much as the spoken data.

It follows that we need other corpora to say something about language learning based on empirical data. For this purpose, this section provides analyses using text transcriptions of what is called *child-directed speech*, namely the speech by an adult directed to a child. The analyses reveal the nature of input children actually obtain and give us some insights into the relation between input stimuli and language learning by children.

5.3.1 The data

There is a large database including text transcriptions of child-adult conversation which is freely available. The database is called *Child Language Data Exchange System*, abbreviated as *CHILDES* (MacWhinney 2000). The database is one component of the *TalkBank* system (<http://talkbank.org/>). It includes a large number of corpora with a variety of languages such as English, German, Spanish, Chinese, Japanese and so on.

Each corpus is written in a format called the CHAT transcription. Here is the example:

```
(171) *CHI: my finger hurts .
      *MOT: what's the matter with your finger # Adam ?
      *CHI: ow # Mommy # ow .
      *CHI: look what I doing to my finger .
      *CHI: I mashing it .
      *URS: but doesn't it hurt # Adam ?
      *CHI: no .
      *URS: why ?
      *CHI: I doing very carefully .
      *URS: don't mash your finger # Adam .
      *CHI: why ?
                                     (from adam37.cha in Brown Corpus [Brown 1973])
```

The sequence surrounded by an asterisk and colon is an id of the speaker of each utterance. Usually the target children is tagged as CHI. Utterances are segmented with words and the letters are in most cases lowered, except some proper nouns (e.g., Adam) and the subjective for of the first person singular pronoun, namely I.

In this section analyses on data from a corpus called *Brown Corpus* (Brown 1973) are provided. The corpus contains transcriptions of conversation between three children and surrounding adults. The children are named with aliases *Adam*, *Eve* and *Sarah*. Data for each child

consists of a subcorpus, referred to as the name of the child (e.g., *Adam Corpus*). Adam Corpus and Sarah Corpus contain utterances by children aged from two to five, but Eve Corpus contains data with a much shorter period. An data overview of the subcorpora is presented in Table 5.29. Each subcorpus is divided into a number of files, each of which is a set of transcription of conversations recorded on a single day. Data recording was done almost on a monthly basis. The recording schedule of Adam Corpus is presented in Table 5.30 (see <http://chilides.psy.cmu.edu/manuals/02english-na.pdf#page=21> for more details). Brown Corpus is used because it contains a reasonable amount of three longitudinal data, though in this study the longitudinal aspect, i.e., some changes during the course of development, is not the main target.

Table 5.29: A data snapshot of Brown Corpus

Child	Age Range	#Files
Adam	2;3–5;2	55
Eve	1;6–2;3	20
Sarah	2;3–5;1	139

Table 5.30: Adam files

File	Age	File	Age	File	Age	File	Age
01	2;03.04	15	2;10.02	29	3;04.18	43	4;01.15
02	2;03.18	16	2;10.16	30	3;05.01	44	4;02.17
03	2;04.03	17	2;10.30	31	3;05.15	45	4;03.09
04	2;04.15	18	2;11.13	32	3;05.29	46	4;04.01
05	2;04.30	19	2;11.28	33	3;06.09	47	4;04.13
06	2;05.12	20	3;00.11	34	3;07.07	48	4;05.11
07	2;06.03	21	3;00.25	35	3;08.01	49	4;06.24
08	2;06.17	22	3;01.09	36	3;08.14	50	4;07.01
09	2;07.01	23	3;01.26	37	3;08.26	51	4;07.29
10	2;07.14	24	3;02.09	38	3;09.16	52	4;09.02
11	2;08.01	25	3;02.21	39	3;10.15	53	4;10.02
12	2;08.16	26	3;03.04	40	3;11.01	54	4;10.23
13	2;09.04	27	3;03.18	41	3;11.14	55	5;02.12
14	2;09.18	28	3;04.01	42	4;00.14		

Note here that the corpus used in the previous study presented in 5.1.2 (Goldberg et al. 2004), namely *Bates Corpus* (Bates et al. 1988), is not used here. The size of the corpus is much smaller than the corpora introduced above, and, though less importantly, it is not a longitudinal corpus, which may have some influence on the interpretation of results if compared to those with longitudinal data.

5.3.2 Method

All the data in CHILDES can be downloaded at the CHILDES website (<http://childes.psy.cmu.edu/>) as text files and hence be freely processed. In this study, therefore, the data are processed using an original script written in a programming language called *Python*.¹⁰⁾ In the script pattern matching is done with regular expression. Specifically, sequences considered to be useful to specify a certain grammatical construction are searched in the corpora, then from the results of the search words appearing at a specific position of the sequences are retrieved, and the frequency distribution of words at the position is calculated. Target data of the search are limited to the speeches by adults, not children, because the research is aimed to know the distributional nature of children's inputs. For the same reason, when the frequency distribution is calculated, the target words are not lemmatized, unlike in the case of large balanced corpora. Children are on their way of language development and hence it is not suitable to utilize such abstracted elements as lemmas.

5.3.3 Analysis

Although it has a reasonable amount of data, the corpus investigated here is not in any sense a large-scaled corpus, so we cannot search such a specific sequence as [... POSS head off] and examine how it behaves in terms of frequency distribution of words appearing at a specific position. From this it follows that sequences suitable to the research on the corpus should have some generality, in the sense that the sequences match a certain number of tokens and that, relatedly, they are used somewhat freely, in not so limited situations.

Given these, the sequences examined here are decided as follows:

- (172) a. [PRONOUN X PRONOUN]
 b. [do you X]

¹⁰⁾<http://www.python.org/>

The former, (172a), is aimed to serve as a surface pattern of the simple transitive construction such as *I love you*. The latter, [do you X], is considered to represent a starting part of an interrogative sentence as seen in *Do you know ... ?*. Both the simple transitive construction and the interrogative construction are two of the most general constructions in English and hence are expected to be used frequently enough in the corpus examined here. In addition, since the *X* slot in the sequences is considered to represent the slot in which a verb appears, examining behavior of them can be anti-arguments against the mainstream verb-centered theories.

The simple transitive construction

Childers & Tomasello (2001) points out that, referring to the results of Jones et al.'s (2000) simulation-based research, the role of pronouns are important for children to learn grammatical structures such as the transitive construction. Jones et al. (2000) finds that there are a number of *pronoun islands* such as *I VERB* and *He Verb*, in the sense that an occurrence of a specific pronoun forms a (syntactic) frame with a slot filled with a verb or other predicates. This specificity is found to be based on the biased frequency of input data, that is, adults speeches the children hear. Therefore, it is hypothesized that such sequences as [PRONOUN ... PRONOUN] in adults' speeches can be useful for children to learn some general syntactic frames. ? actually examine the hypothesis by experiments using frames such as *He's [VERB]-ing it* and find that the pronoun-full frames are significantly more effective to learn syntactic structures than those with lexical nouns. With the findings in mind, in this research, the sequence [PRONOUN X PRONOUN] is investigated as one functioning as a surface pattern of the simple transitive construction. Specifically, the data are searched with a regular expression presented below:

```
(173) \b(?:I|he|she|we|they) (\S+?) (?:me|you|him|her|it|them|us)
      [.!?,;:]
```

This means a sequence starts with either *I*, *he*, *she*, *we* or *they* followed by any single word and ends with any punctuation marks, preceded by pronouns *me*, *you*, *him*, *her*, *it*, *them* or *us*.

Results for the three children's data are shown in Table 5.31-5.33 and Figure 5.29-5.31. All the tables and graphs clearly show that the frequency distribution of words at *X* position obey Zipf's law, though the frequently observed words are different from child to child. Examples found in Adam Corpus are presented below (underline added):

(174) a. *MOT: I like it . (adam01)

- b. *URS: Adam # may I try it ? (adam26)¹¹⁾
 c. *MOT: you mean did she take them ? (adam40)

Table 5.31: Adam: X
in [PRONOUN X
PRONOUN]

rank	X	freq.
1	see	27
2	call	8
3	like	5
4	tie	4
5	help	3
5	show	3
5	tell	3
8	catch	2
8	have	2
8	needs	2
8	fit	2
8	saw	2
8	do	2
8	put	2
8	heard	2
8	dropped	2
⋮		
total		95

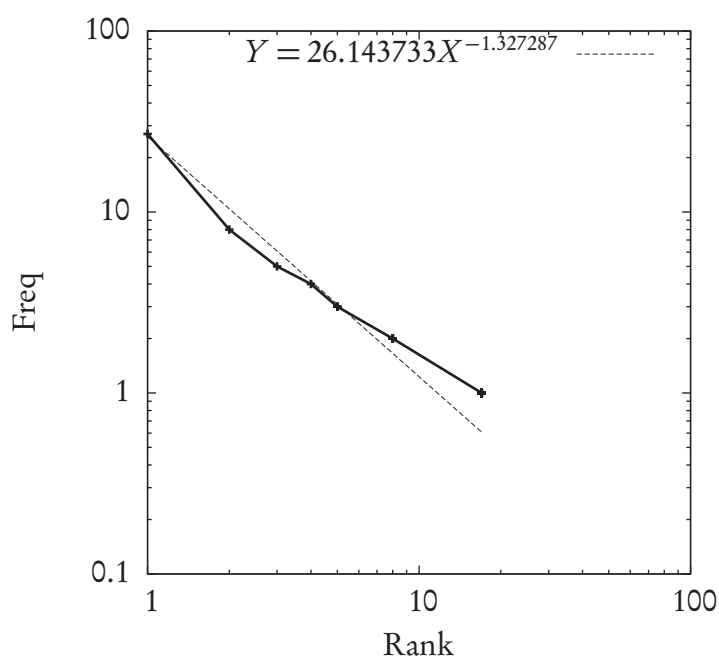


Figure 5.29: Adam: X in [PRONOUN X PRONOUN]

Incidentally, the same distributional tendency is found in the *PRONOUN-PRONOUN* pairs seen in the matched sequences with the pattern. In the case of Adam the frequency of the pairs is distributed as shown in Table 5.34. As exemplified in (174a) and (174b), the pair *I-it* is the most frequent, whose frequency is almost twice as high as that of the second-most, *I-you*.

The interrogative construction with *do you*

The next sequence is [do you X], which is considered to be a starting part of an interrogative sentence. The regular expression used here is the following:

¹¹⁾URS represents a name of an investigator, Ursula Bellugi.

Table 5.32: Eve: X
in [PRONOUN X
PRONOUN]

rank	X	freq.
1	see	26
2	get	4
2	hear	4
4	know	3
5	help	2
5	buy	2
5	have	2
5	stir	2
5	like	2
	\vdots	
total		67

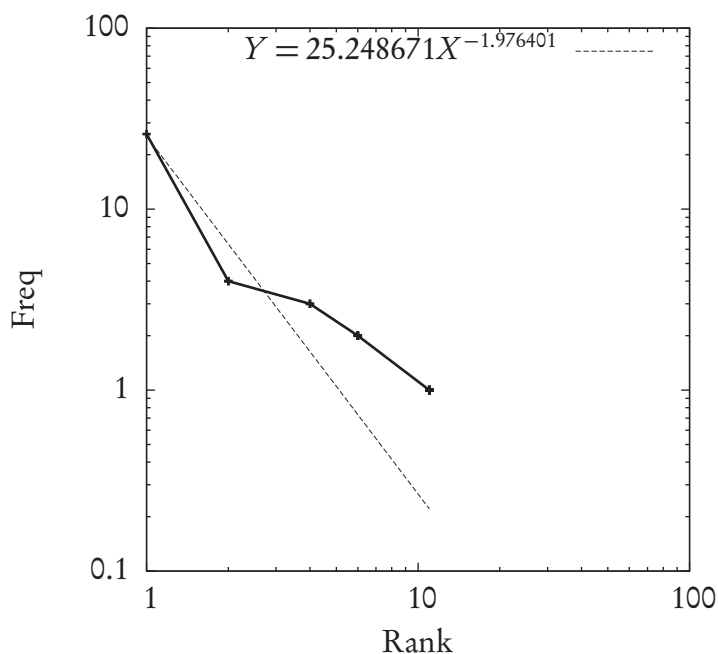


Figure 5.30: Eve: X in [PRONOUN X PRONOUN]

Table 5.33: Sarah: X
in [PRONOUN X
PRONOUN]

rank	X	freq.
1	know	25
2	see	13
3	bite	10
4	love	9
5	like	8
6	got	7
7	tell	6
8	give	5
8	put	5
10	saw	4
10	told	4
10	do	4
10	cut	4
	\vdots	
total		197

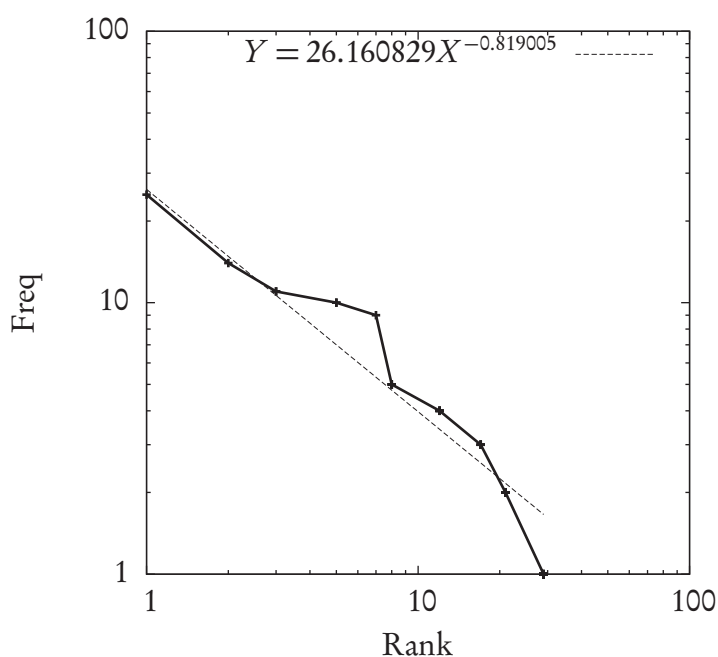


Figure 5.31: Sarah: X in [PRONOUN X PRONOUN]

Table 5.34: Adam: *PRONOUN-PRONOUN* in [PRONOUN X PRONOUN]

rank	pro-pro	freq.
1	I-it	32
2	I-you	18
3	he-it	11
4	she-you	8
5	he-you	7
6	we-it	6
7	I-them	4
8	they-them	2
8	they-you	2
8	she-it	2
11	she-them	1
11	they-it	1
11	they-him	1

(175) \bdo you (\S+)\b

This means a sequence starting with *do you*, followed by any single word.

The results are shown in Table 5.35-5.37 and Figure 5.32-5.34. The *A* value of Adam's and the *B* value of Sarah's are somewhat low, but not to a large extent. The other values almost fit the model. Therefore it can be concluded that words at *X* slot in the sequence also shows Zipfian rank-freq distribution. Examples from Eve Corpus are presented as below (underline added):

- (176) a. *MOT: what do you want me to do with it ? (eve01)
 b. *FAT: do you want something ? (eve07)
 c. *MOT: do you see Eve in there ? (eve15)

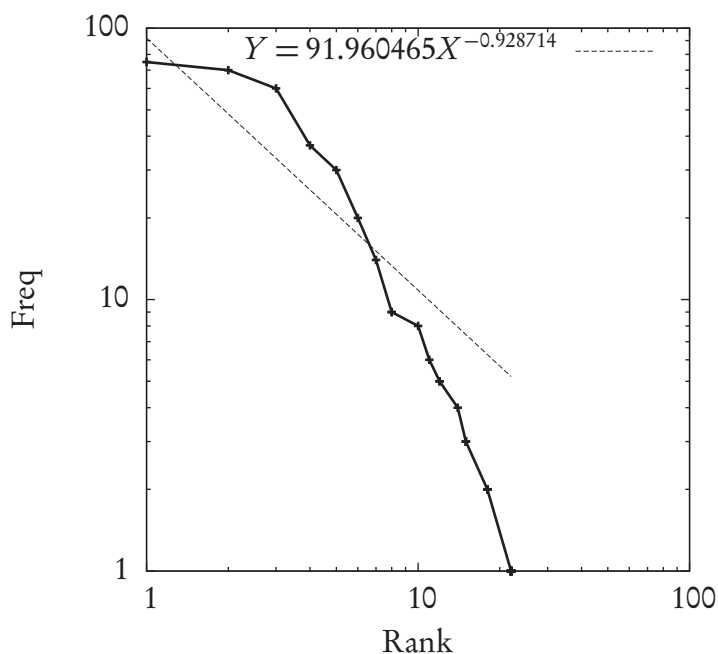
Many examples (59 out of 173) accompany the interrogative *what*, as seen in (176a), which may mean the sequence [what do you ...] also functions as a surface pattern.

5.4 Concluding remarks

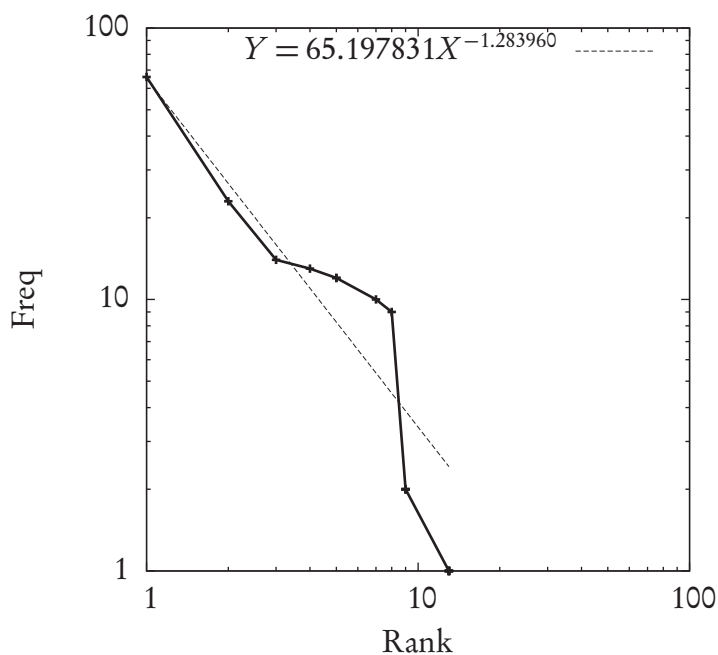
This chapter provides a number of quantitative analyses on sequences which presumably represent major constructions in English including those investigated in the last chapter, namely

Table 5.35: Adam: X in [do you X]

rank	X	freq.
1	want	75
2	have	70
3	know	60
4	think	37
5	see	30
6	like	20
7	remember	14
8	need	9
8	do	9
10	say	8
	\vdots	
total		387

Figure 5.32: Adam: X in [X *me* ARTICLE]Table 5.36: Eve: X in [do you X]

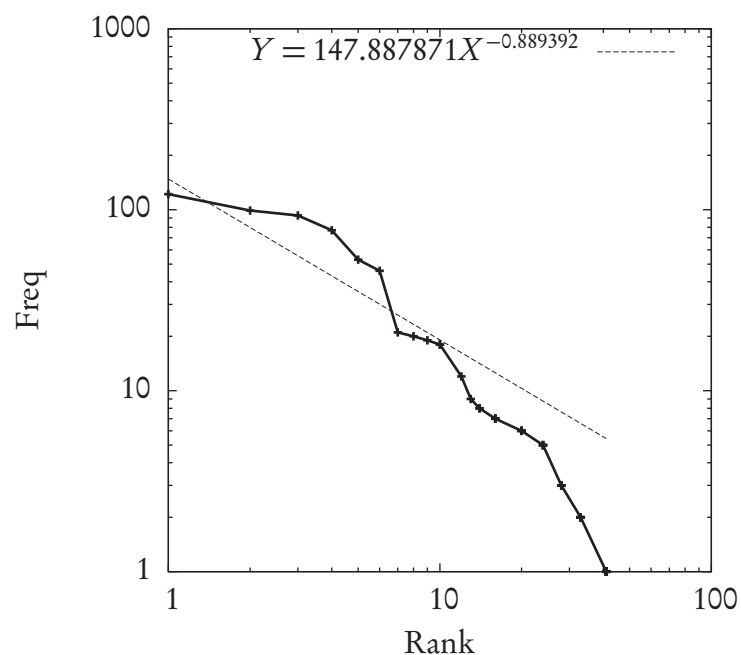
rank	X	freq.
1	want	66
2	have	23
3	like	14
4	see	13
5	do	12
5	know	12
7	say	10
8	think	9
9	remember	2
9	hear	2
9	not	2
9	mean	2
	\vdots	
total		173

Figure 5.33: Eve: X in [X *me* ARTICLE]

the ditransitive construction, the resultative construction, the caused-motion construction and the *way* construction. Though the analyses yielded somewhat mixed results, the research conducted in this chapter is generally considered successful. However, there are some prob-

Table 5.37: Sarah: *X* in [do you *X*]

rank	<i>X</i>	freq.
1	want	122
2	think	99
3	like	93
4	know	77
5	have	53
6	do	46
7	call	21
8	say	20
9	see	19
10	remember	18
10	wanna	18
⋮		
total		755

Figure 5.34: Sarah: *X* in [*X me ARTICLE*]

lems mostly on the methodological aspect. Below a brief summary of this chapter and the problems are presented.

5.4.1 Brief summary

The results of the research presented in this chapter show that the sequence considered to function as surface patterns of some constructions actually contain slots at which items are distributed in the way showing Zipfian rank-freq distribution, as hypothesized in 151. This is confirmed mainly by the results investigating sequences related to English major constructions, such as [... PRONOUN ARTICLE ...] for the ditransitive construction and [... to death] for the resultative construction. Many can be said successful, but the distribution of word at *X* position in [*X POSS head off*], a candidate of a surface pattern of the resultative construction, is not considered to be Zipfian. That would be because of the ambiguity of the sequence, and therefore it would probably not function as a surface pattern of the construction. The distributional tendency is examined not only for the data from large-scale corpus mainly containing written language, but also for spoken utterances which are directed to children. In both types of data expected results are obtained

5.4.2 Limitations and problems

Mainly methodologically, a couple of limitations and problems can be pointed out for the research presented in this chapter. They are expressed as forms of two *why*-questions:

- (177) a. Why single words?
 b. Why single positions?

Why single words?

The first question is about the items appearing at a specific position, in many cases represented as *X* in the research presented above. For all the sequences analyzed in this chapter the items whose frequency distribution is examined are single words, as opposed to word sequences. There are, however, many cases where a sequence with two or more words appear at a specific position of a pattern with slots. For example, the position *X* in the sequence [PRONOUN *X* PRONOUN] can be filled with two-word sequences such as *just read* (178a), *can do* (178b) and *played with* (178c), as exemplified below (from Adam Corpus; underline added):

- (178) a. *MOT I just read them . (adam04)
 b. *URS: ask Cromer if he can do it . (adam17)
 c. *MOT: he played with you ? (adam26)

Such multi-word sequences can, just like words, also be recognized as elements appearing at a certain slot. If so, we have to incorporate those sequences into the targets of distributional analysis.

Here arises a problem: sequence of how many maximal words are allowed to participate as members of elements appearing at a slot? It may be two, but may also be three and more. There is no clear criterion to set the maximal number of words regarded as elements filling a slot. In addition, more problematically, in the cases where the slot is located at the beginning or the end, as opposed to the middle, of a sequence, the boundary of sequence itself is drastically blurred, which may probably cause the results messy. For example, the *X* position of [*X* me ARTICLE] can be expanded to include two-word sequences, but the results can be hardly interpretable. Table 5.38 shows the frequent two-word sequences appearing just before [me a/an]. They can be classified into several groups: those with “punctuation mark+verb,” “subject+verb,” “conjunctive+verb,” and so on, but cannot be generalized into one single distributional characterization.

Table 5.38: X-Y in [X Y me a/an] on COCA

rank	X-Y	freq.
1	,-give	839
2	it-take	770
3	he-give	720
4	and-give	656
5	to-give	653
6	.-give	551
7	she-give	522
8	you-give	500
9	it-give	497
10	“-give	482

For now there is no clear solution to this problem, but a promising hypothesis is as follows: items appearing just before and after, but not between, some fixed sequence should be single-segment elements such as words; those occurring between fixed elements, in contrast, can span multiple segments. This is based on the fact that cross-linguistically, there are many expressions or syntactic frames with an internal slot between fixed edges. The typical examples are German *Satzklammer* (also known as the *sentence bracket*) and the English verb-particle construction, as exemplified below:

- (179) a. Ich kann Hans sehen.
 I can.1stSg Hans see.inf
 “I can see Hans.”
- b. Wenn ich meine Tante besuchte, backte sie für mich Kuchen.
 when I my.fem aunt visit.past bake.past she for me.acc cake
 “When I visited my aunt, she baked a cake for me.”
- c. You should check it out.
- d. I don’t want to give it up.

In addition to them, similar structures can be found elsewhere in English such as in the sequence composed of an article and noun (e.g., *the tall man*) and in that with a preposition and a noun (e.g., *to some extent*). Examples of such structure are frequently found probably because of their, say, *cognitive utility*: if elements at the edges of a sequence are fixed we can easily find the sequence to be a pattern with an internal slot, which can allow itself to have

multi-length slot fillers.

Why single positions?

In addition to the *single word* problem posed above, what can be called *single position problem* is also problematic. For all the sequences examined in this chapter, only one position at each sequence is analyzed. There is, however, a possibility that elements whose frequency distribution obeys Zipf's law are not those appearing at a single position, but those spanning multiple slots. For example, the sequence [... *me* ARTICLE] can be analyzed in terms of pairs of items appearing before *me* and after *ARTICLE*, just like a research by Yoshikawa (2010) presented in the previous chapter (4.2.2). The sequence [... *me* a/an ...] matches pairs on COCA shown in Table 6.3, whose frequency distribution, though, cannot be seen as obeying Zipf's law.

Table 5.39: X - Y in [X *me* a/an Y] on COCA

rank	X - Y	freq.
1	give-break	545
2	do-favor	495
3	give-chance	449
4	give-little	308
5	tell-little	272
6	give-call	265
7	give-sense	251
8	take-while	244
9	take-long	218
10	tell-story	198

This problem should probably be solved on a case-by-case basis in the sense that sequences with multiple slots are examined in terms both of the items appearing at each single slot and of those spanning multiple slots. If the frequency distribution of items appearing at multiple slots in a sequence are found to obey Zipf's law, we should recognize the sequence as a pattern with discontinuous slots.

Chapter 6

Concluding remarks

In this dissertation a theory of grammatical constructions named *Exemplar-based Construction Grammar* is presented and based on the assumptions and methodology a number of case studies are conducted to show the validity of the theory. The main targets of analysis are English major constructions such as the ditransitive construction and resultative construction.

6.1 Exemplar-based characterizations of constructions

In Chapter 3 the theoretical conceptions of the framework presented in this dissertation are described in great detail. The chapter presents, for example, the definition of construction under the exemplar-based framework, which reinterprets constructions as *phenomena*, as opposed to *entities*, derived from association of exemplars with a partially-lexically fixed sequence called a *surface pattern* such as [She ... me a question] as in *He kicked me a question*. This is an application of the widely-shared view of cognitive processing in what is called *exemplar theory*, as introduced in Chapter 2. Constructions reinterpreted as phenomena are called *construction effects*.

In addition to the theoretical aspect, Chapter 3 also includes some remarks on the methodological aspects of Exemplar-based Construction Grammar (EBCG). In order to verify its assumptions EBCG utilizes the surface patterns which are considered effective to specify a certain constructions. For example the sequence [... will ... me a ...] is seen to be usable to collect examples of the ditransitive construction as in *You will give me a kiss*. At the same time, however, the sequence also matches those instantiating constructions other than the ditransitive, such as *You will miss me a lot*. In that case we have to find other sources of the non-targeted constructions, that is, a set of exemplars associated with some other surface pat-

terns, such as ... *miss* ... *a lot*. If such competing patterns are not in an *inclusive* relation as seen in, for example, [A B ...] and [A ...], they are considered to function as *evokers* of constructions independently. In this way, the theory is partially confirmed if of all the examples matching a (set of) surface pattern(s), those instantiating the target construction (e.g., the ditransitive construction) are successfully collected and those not instantiating it are excluded.

In Chapter 4 English four constructions analyzed by a number of previous studies, the ditransitive, resultative, caused-motion and *way* construction, are described under the exemplar-based framework. Let us take the ditransitive construction as an example. Based on a previous study by Du Bois (2003), the sequence [VERB PRONOUN ARTICLE NOUN] is found to be a leading candidate for a surface pattern, but since it still contains two open-class category, *VERB* and *NOUN*, more lexically unspecified sequence [... PRONOUN ARTICLE ...] is also presented as a candidate.

As a case study using the former candidate as a query sequence of a corpus search, Yoshikawa's (2010) research is presented, in which some combinations between certain verbs appearing at *VERB* position and nouns at *NOUN* position of the sequence have highly preferred associations and forming a kind of meaning clusters, which, in turn, can be seen as the real source of *constructional polysemy* discussed by Goldberg (1995).

Since the case study focuses on typical pairs of certain verbs and nouns, cases with a non-typical verb or noun cannot be explained based only on the Yoshikawa's (2010) findings. Moreover, although the candidate for a surface pattern includes pronoun at the position where an indirect object is expected to fill in, there are also cases in which a lexical noun phrase functions as an indirect object. In order to solve those problems, some additional explanations are offered, mainly by showing even in those problematic cases there would be some other cues which successfully navigate us to the target construction. However, ditransitive sentences whose indirect object is *proper* noun cannot be explained.

6.2 Exemplars and Zipf's law

In Chapter 5 a well-known frequency distribution often found in the distribution of linguistic elements is applied to the surface-pattern-based analyses of grammatical constructions. The distribution is called *Zipfian distribution*, named after the discoverer Jorge Kingsley Zipf (Zipf 1935, 1949). The distribution is one seen between frequency and rank of the type of elements such as words, phonemes and letters in a certain distribution space such as a set of texts

in a corpus, in which the two variables, frequency and rank, are inversely logarithmically-proportional to each other. Specifically, the frequency of words at the rank R , S_R , should be S_R/S_1 if the frequency distribution obeys Zipf law.

Zipf's law is not only a mere regularity found in frequency distribution, but also related to the nature of our quantity perception. We human tend to perceive quantity in a logarithmic way, represented in what is called *Weber-Fechner law* (Varshney & Sun 2013). From this it follows that the discretely logarithmic/exponential quantity distribution being the optimal condition for us to recognize the difference between each items at each rank. In addition, perhaps related to this, the Zipfian-like skewed input frequency is known to be beneficial for both children and adult in learning grammatical constructions of first and second language (Casenhiser & Goldberg 2005; Goldberg et al. 2004; Ellis & Ferreira-Junior 2009).

The mathematical properties of Zipf's law is also investigated in order to prepare for analyses of corpus-based research. Based on the mathematics of Zipf's law, the actual frequency distributions of words at certain position of several (candidates for) surface patterns are investigated. Data are mainly retrieved from a large-scale balanced corpus named *Corpus of Contemporary American English* (COCA: Davies 2008-), but a small scale corpus with child-adult conversations are also used for the purpose of exploring the nature of children's inputs in the course of language development. The results are generally as expected, though there remain some problems.

6.3 Future issues

As mentioned in the previous chapters, there are some problems yet to be solved. Of them the most problematic is, as discussed in 3.7.2, the lack of production model. Linguistic phenomena related to grammatical constructions can probably not be fully explained only with the comprehension model. We have to explain where and how such and such expressions come from.

In addition to this, for now the following two problems can be pointed out:

- the problem of unit detection
- the problem (of the use) of statistics

The first problem is one related to that discussed in 5.4.2. Linguistic sequences we actually

experience are continuous stream and hence we have to detect boundaries of a unit to extract it. For now the current framework does not provide any principled procedures to achieve unit detection.

The second problem would be somewhat straight forward. The quantitative analyses presented in Chapter 5 have some statistic character, but the numerical data actually used are only the raw frequencies of words. In the literatures of corpus linguistics it is often pointed out that the mere raw frequency is not adequate, or in some cases implausible (e.g., Stefanowitsch & Gries 2003), which means that some kind of statistics are needed to complement the possible deficiency of using raw frequency.

Bibliography

- Arnon, I. & N. Snider. 2010. More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62(1):67 – 82.
- Baddeley, A. D. & G. Hitch. 1974. Working memory. In G. H. Bower, ed., *The psychology of learning and motivation*, vol. 8, pp. 47–89. New York: Academic Press.
- Barnbrook, G. 1996. *Language and computers: A practical introduction to the computer analysis of language*. Edinburgh: Edinburgh University Press.
- Bates, E., I. Bretherton, & L. S. Snyder. 1988. *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge: Cambridge University Press.
- Bird, S., E. Klein, & E. Loper. 2009. *Natural language processing with Python*. Beijing; Cambridge, MA.: O'reilly.
- Boas, H. C. 2003. *A constructional approach to resultatives*. Stanford: CSLI Publications.
- Boas, H. C. & I. A. Sag, eds. 2012. *Sign-based construction grammar*. Stanford: CSLI Publications.
- Bod, R. 2009. From Exemplar to Grammar: A Probabilistic Analogy-Based Model of Language Learning. *Cognitive Science* 33(5):752–793.
- Borensztajn, G. 2011. *The neural basis of structure in language: Bridging the gap between symbolic and connectionist models of language processing*. Amsterdam: Institute for Logic, Language and Computation, Universiteit van Amsterdam.
- Borensztajn, G. & W. Zuidema. 2011. Episodic grammar: A computational model of the interaction between episodic and semantic memory in language processing. In *Proceedings of the 33th Annual Conference of the Cognitive Science Society*, pp. 507–512.

- Brown, R. 1973. *A first language: The early stages*. Cambridge, MA.: Harvard University Press.
- Bybee, J. 1995. Regular morphology and the lexicon. *Language and cognitive processes* 10(5):425–455.
- Bybee, J. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Carrier, J. & J. H. Randall. 1992. The argument structure and syntactic structure of resultatives. *Linguistic inquiry* 23(2):173–234.
- Casenhiser, D. & A. E. Goldberg. 2005. Fast mapping between a phrasal form and meaning. *Developmental Science* 8(6):500–508.
- Chafe, W. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Childers, J. B. & M. Tomasello. 2001. The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology* 37(6):739–748.
- Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. 1980. *Rules and representations*. New York: Columbia University Press.
- Chomsky, N. 1993. *Lectures on government and binding: The Pisa lectures*. Berlin; New York: Mouton de Gruyter, 7th edn.
- Colleman, T. & B. De Clerck. 2009. 'Caused motion'?: The semantics of the English to-dative and the Dutch aan-dative. *Cognitive Linguistics* 20(1):5–42.
- Croft, W. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford; New York: Oxford University Press.
- Croft, W. 2003. Lexical rules vs. constructions: A false dichotomy. In H. Cuyckens, R. Dirven, & K.-U. Panther, eds., *Motivation in language*, pp. 49–68. Amsterdam; Philadelphia: John Benjamins.
- Croft, W. 2005. Logical and typological arguments for Radical Construction Grammar. In M. Fried & J.-O. Östman, eds., *Construction Grammars: Cognitive grounding and theoretical extensions*, pp. 273–314. Amsterdam; Philadelphia: John Benjamins.

- Davies, M. 2004-. BYU-BNC: The British National Corpus (Based on the British National Corpus from Oxford University Press). Available online at <http://corpus.byu.edu/bnc/>.
- Davies, M. 2008-. The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present. Available online at <http://corpus.byu.edu/coca/>.
- Davies, M. 2011-. Google Books (American English) Corpus (155 billion words, 1810-2009). Available online at <http://googlebooks.byu.edu/>.
- Dietrich, R., W. Klein, & C. Noyau. 1995. *The acquisition of temporality in a second language*, vol. 7. John Benjamins Publishing.
- Du Bois, J. W. 2003. Argument structure. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby, eds., *Preferred argument structure: Grammar as architecture for function*, pp. 11–60. Amsterdam; Philadelphia: John Benjamins.
- Du Bois, J. W., W. L. Chafe, C. Meyer, & S. A. Thompson. 2000. *Santa Barbara corpus of spoken American English, Part 1*. Philadelphia: Linguistic Data Consortium.
- Du Bois, J. W., W. L. Chafe, C. Meyer, S. A. Thompson, & N. Martey. 2003. *Santa Barbara corpus of spoken American English, Part 2*. Philadelphia: Linguistic Data Consortium.
- Du Bois, J. W. & R. Englebretson. 2004. *Santa Barbara corpus of spoken American English, Part 3*. Philadelphia: Linguistic Data Consortium.
- Du Bois, J. W. & R. Englebretson. 2005. *Santa Barbara corpus of spoken American English, Part 4*. Philadelphia: Linguistic Data Consortium.
- Eichenbaum, H., P. Dudchenko, E. Wood, & M. Shapiro. 1999. The hippocampus, memory, review and place cells: Is it spatial memory or a memory space? *Neuron* 23:209–226.
- Ellis, N. C. & F. Ferreira-Junior. 2009. Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics* 7(1):187–220.
- Feldweg, H. 1991. *The European science foundation second language database*. Nijmegen: Max-Planck-Institute for Psycholinguistics.
- Fillmore, C., P. Kay, & M. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language* 64(3):501–538.

- Ganter, B., G. Stumme, & R. Wille. 2005. *Formal Concept Analysis: Foundations and applications*, vol. 3626. Berlin: Springer.
- Ganter, B., R. Wille, & C. Franzke. 1997. *Formal concept analysis: Mathematical foundations*. Berlin: Springer.
- Gerken, L., R. Wilson, & W. Lewis. 2005. Infants can use distributional cues to form syntactic categories. *Journal of child language* 32(2):249–268.
- Givón, T. 2002. *Bio-linguistics: The Santa Barbara lectures*. Amsterdam; Philadelphia: John Benjamins.
- Goldberg, A. E. 1991. It can ' t go down the chimney up: Paths and the English resultative. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, vol. 17, pp. 368–378.
- Goldberg, A. E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago; London: University of Chicago Press.
- Goldberg, A. E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, A. E., D. M. Casenhiser, & N. Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15(3):289–316.
- Goldberg, A. E. & R. Jackendoff. 2004. The English resultative as a family of constructions. *Language* pp. 532–568.
- Goldinger, S. D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(5):1166–1183.
- Gries, S. T. 2003. Towards a corpus-based identification of prototypical instances of constructions. *Annual review of cognitive linguistics* 1(1):1–27.
- Gries, S. T., B. Hampe, & D. Schönefeld. 2005. Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16(4):635–676.
- Gropen, J., S. Pinker, M. Hollander, R. Goldberg, & R. Wilson. 1989. The learnability and acquisition of the dative alternation in English. *Language* pp. 203–257.

- Hawkins, J. & S. Blakeslee. 2004. *On intelligence*. New York: Macmillan.
- Hintzman, D. L. 1984. MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers* 16(2):96–101.
- Hintzman, D. L. 1986. “Schema abstraction” in a multiple-trace memory model. *Psychological Review* 93(4):411–428.
- Iwata, S. 2008. *Locative alternation: A lexical-constructional approach*. Amsterdam; Philadelphia: John Benjamins.
- Jackendoff, R. 1977. *\bar{X} syntax: a study of phrase structure*. Cambridge, MA.: the MIT Press.
- Jackendoff, R. 1990. *Semantic structures*, vol. 18. Cambridge, MA.: MIT Press.
- Jackson, A. & J. Morton. 1984. Facilitation of auditory word recognition. *Memory & Cognition* 12(6):568–574.
- Johnson, K. 1997. Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix, eds., *Talker variability in speech processing*, pp. 145–165. San Diego: Academic Press.
- Johnson, K. 2006. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34:485–499.
- Jones, G., F. Gobet, & J. M. Pine. 2000. A process model of children’s early verb use. In *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society*, pp. 723–728. Erlbaum.
- Jurafsky, D. S. 1992. *An on-line computational model of human sentence interpretation: A theory of the representation and use of linguistic knowledge*. Ph.D. thesis, University of California, Berkeley.
- Kay, P. & C. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the *What’s X doing Y?* construction. *Language* 75(1):1–33.
- Kruschke, J. K. 1992. ALCOVE: An exemplar-based connectionist model of category learning. *Psychological review* 99(1):22–44.
- Kunihiro, T. 1967. *Kozo-teki imiron: Nichi-Ei ryogo taisho kenkyu* [*Structural semantics: A contrastive study of English and Japanese*]. Tokyo: Sanseido.

- Kuno, S., K. Takami, K. Courtenay, & N. Decker. 2004. The way construction and unergativity. In S. Kuno & K. Takami, eds., *Functional constraints in grammar: On the unergative-unaccusative distinction*, pp. 67–104. Amsterdam: John Benjamins.
- Kuroda, K. 2000. *Foundations of Pattern Matching Analysis: A new method for the cognitively realistic description of natural language syntax*. Ph.D. thesis, School of Human and Environmental Studies, Kyoto University.
- Kuroda, K. 2009. Pattern lattice as a model for linguistic knowledge and performance. In *Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation*, pp. 278–287. Hong Kong.
- Kuromiya, K. 2010. Where does the Motion sense of the caused-motion construction come from? In *Proceedings of the Annual Meetings of the Japanese Cognitive Linguistics Association*, vol. 10, pp. 405–415.
- Lakoff, G. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago press.
- Langacker, R. 1987. *Foundations of cognitive grammar Vol 1.: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, R. 2000. A dynamic usage-based model. In M. Barlow & S. Kemmer, eds., *Usage-based models of language*, pp. 1–63. Stanford: CSLI Publications.
- Langacker, R. 2009. A dynamic view of usage and language acquisition. *Cognitive Linguistics* 20(3):627–640.
- Langacker, R. W. 1991. *Foundations of Cognitive Grammar: Descriptive application. Volume 2*. Stanford: Stanford University Press. ISBN 9780804738521.
- Levin, B. & M. Rappaport Hovav. 1991. Wiping the slate clean: A lexical semantic exploration. *Cognition* 41(1):123–151.
- Lieven, E. V., J. M. Pine, & G. Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of child language* 24(01):187–219.
- MacWhinney, B. 2000. *The CHILDES project: Tools for analyzing talk*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

- Marcus, G. F., S. Vijayan, S. B. Rao, & P. M. Vishton. 1999. Rule learning by seven-month-old infants. *Science* 283(5398):77–80.
- Markman, E. M. 1990. Constraints children place on word meanings. *Cognitive Science* 14(1):57–77.
- McClelland, J., D. Rumelhart, & P. R. Group. 1986. *Parallel Distributed Processing*, vol. 1. Cambridge, MA.: MIT Press. ISBN 9780262680530.
- Medin, D. L. & M. M. Schaffer. 1978. Context theory of classification learning. *Psychological review* 85(3):207–238.
- Michel, J.-B., Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, T. G. B. Team, J. P. Pickett, D. Hoiberger, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M. A. Nowak, & E. L. Aiden. 2011. Quantitative analysis of culture using millions of digitized books. *Science* 331(6014):176–182.
- Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review* 63(2):81–97.
- Mintz, T. H. 2002. Category induction from distributional cues in an artificial language. *Memory & Cognition* 30(5):678–686.
- Mintz, T. H. 2003. Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90(1):91–117.
- Mukherjee, J. 2005. *English ditransitive verbs: Aspects of theory, description and a usage-based Model*, vol. 53. Amsterdam: Rodopi.
- Nakatani, K. 2007. Bun-shori-strategy to-iu shiten kara mi-ta kekka-kobun no ruikei-ron[A typology of the resultative construction in terms of the sentence-processing strategy]. In N. Ono, ed., *Kekka-kobun kenkyu no shin-shiten [New perspectives of the study of the resultative construction]*, pp. 289–317. Tokyo: Hituzi-shobo.
- Newman, J. 1996. *Give: A cognitive linguistic study*. Berlin: Walter de Gruyter.
- Nosofsky, R. M. 1984. Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10(1):104.

- Nosofsky, R. M. 1986. Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General* 115(1):39–57.
- Palmeri, T. J., S. D. Goldinger, & D. B. Pisoni. 1993. Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:309.
- Perdue, C., ed. 1993. *Adult language acquisition. Vol 1: Field methods*. Cambridge: Cambridge University Press.
- Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In P. Hopper & J. Bybee, eds., *Frequency and the emergence of linguistic structure*, pp. 137–157. Amsterdam; Philadelphia: John Benjamins.
- Pollard, C. & I. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- Port, R. F. 2003. Meter and speech. *Journal of Phonetics* 31(3):599–611.
- Port, R. F. 2007a. How are words stored in memory?: Beyond phones and phonemes. *New Ideas in Psychology* 25(2):143–170.
- Port, R. F. 2007b. The problem of speech patterns in time. In M. Gaskell, ed., *The Oxford handbook of psycholinguistics*, pp. 503–514. Oxford; New York: Oxford University Press.
- Port, R. F. 2010. Rich memory and distributed phonology. *Language Sciences* 5:43–55.
- Roberts, L. D. 2004. The relation of children’s early word acquisition to abduction. *Foundations of science* 9(3):307–320.
- Rumelhart, D., J. McClelland, & P. R. Group. 1986. *Parallel Distributed Processing*, vol. 1. Cambridge, MA.: MIT Press. ISBN 9780262680530.
- Saffran, J. R., R. N. Aslin, & E. L. Newport. 1996a. Statistical cues in language acquisition: Word segmentation by infants. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, pp. 376–380.
- Saffran, J. R., E. L. Newport, & R. N. Aslin. 1996b. Word segmentation: The role of distributional cues. *Journal of memory and language* 35(4):606–621.

- Sag, I. A., H. C. Boas, & P. Kay. 2012. Introducing sign-based construction grammar pp. 2–29.
- Simpson, J. 1983. Resultatives. In L. Levin, M. Rappaport, & A. Zaenen, eds., *Papers in Lexical-Functional Grammar*, pp. 143–157. Bloomington: Indiana University Linguistics Club.
- Sinclair, J. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press. ISBN 0194371441.
- Stefanowitsch, A. 2008. Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics* 19(3):513–531.
- Stefanowitsch, A. & S. T. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International journal of corpus linguistics* 8(2):209–243.
- Stefanowitsch, A. & S. T. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1):1–43.
- Sun, J. Z., G. I. Wang, V. K. Goyal, & L. R. Varshney. 2012. A framework for bayesian optimality of psychophysical laws. *Journal of Mathematical Psychology* 56(6):495–501.
- Taylor, J. R. 2012. *The mental corpus: How language is represented in the mind*. Oxford: Oxford University Press.
- Tomasello, M. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA.: Harvard University Press.
- Tulving, E. 1972. Episodic and semantic memory. In E. Tulving & W. Donaldson, eds., *Organization of memory*, pp. 381–402. New York: Academic Press.
- Varshney, L. R. & J. Z. Sun. 2013. Why do we perceive logarithmically? *Significance* 10(1):28–31.
- Vihman, M. M. & T. Keren-Portnoy. 2013. *The emergence of phonology: Whole-word approaches and cross-linguistic evidence*. Cambridge: Cambridge University Press. ISBN 9780521762342.
- Wray, A. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

- Wray, A. 2007. 'Needs only' analysis in linguistic ontogeny and phylogeny. In C. Lyon, C. L. Nehaniv, & A. Cangelosi, eds., *Emergence of communication and language*, pp. 53–70. London: Springer.
- Wray, A. & G. W. Grace. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117:543–578.
- Yoshikawa, M. 2009. The “beginning paradox”: An empirical and conceptual problem in the current Usage-based Model of language. *Colloquia* 30:55–72.
- Yoshikawa, M. 2010. “Kobun no tagi” sai-kou: Jirei-kiban-kobun-riron ni muke-te [Re-thinking “constructional polysemy”: Toward an exemplar-based theory of grammatical constructions]. In *Proceedings of 10th Annual Meeting of the Japanese Cognitive Linguistics Association*, pp. 449–459.
- Zipf, G. K. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Boston: Houghton Mifflin.
- Zipf, G. K. 1949. *Human behavior and the principle of least effort*. Boston: Addison-Wesley.