



Universiteit
Leiden
The Netherlands

Constructions emerging : a usage-based model of the acquisition of grammar

Beekhuizen, B.F.

Citation

Beekhuizen, B. F. (2015, September 22). *Constructions emerging : a usage-based model of the acquisition of grammar*. *LOT dissertation series*. LOT, Utrecht. Retrieved from <https://hdl.handle.net/1887/35460>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/35460>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/35460> holds various files of this Leiden University dissertation

Author: Beekhuizen, Barend

Title: Constructions emerging : a usage-based model of the acquisition of grammar

Issue Date: 2015-09-22

Constructions Emerging
A Usage-Based Model
of the Acquisition of Grammar

The research reported here was supported by NWO grant 322.70.001

Published by

LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 5775
e-mail: lot@uu.nl
<http://www.lotschool.nl>

Cover illustration: #9f4c62

ISBN: 978-94-6093-183-3
NUR: 616

Copyright © 2015 Barend F. Beekhuizen. All rights reserved.

Constructions Emerging
A Usage-Based Model
of the Acquisition of Grammar

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 22 september 2015
klokke 13.45 uur

door

Barend F. Beekhuizen

geboren 7 januari 1987
te 's Gravenhage, Nederland

Promotores: Prof. Dr. Arie Verhagen
Prof. Dr. Rens Bod (University of Amsterdam)

Promotiecommissie: Prof. Dr. Roberta D'Alessandro
Dr. Afra Alishahi (Tilburg University)
Prof. Dr. Ewa Dąbrowska (Northumbria University)

There is a crack in everything
That's how the light gets in.
Leonard Cohen, "Anthem"

Little feet take small steps
Bloom (1991, 11)

Contents

Acknowledgements	xi
1 Introduction	1
1.1 Early grammar	3
1.2 Theoretical background	4
1.3 Computational cognitive modeling	5
1.4 Goals of this research	6
1.4.1 Providing a comprehensive model	7
1.4.2 The conception of learning	7
1.4.3 Starting small	8
1.4.4 Naturalism in meaning	8
1.5 A note on notation	9
1.6 Overview of the dissertation	10
2 A usage-based conception of language acquisition	11
2.1 Usage-based linguistics and language acquisition	12
2.1.1 Constructions and the constructicon	14
2.1.2 Producing and understanding an utterance	17
2.1.3 Acquiring a grammar	18
2.2 Theoretical issues with the usage-based perspective	21
2.2.1 Representational metaphors: blocks and streams	21
2.2.2 Mechanisms operating on early representations	22
2.2.3 Gradualism and simultaneity in learning	24
2.3 Desiderata for a usage-based model of language acquisition	27
2.3.1 D1: Explicitness	28
2.3.2 D2: Comprehensiveness	28
2.3.3 D3: Simultaneity	28
2.3.4 D4: Cognitive realism in representations	29
2.3.5 D5: Cognitive realism in processes	30
2.3.6 D6: Cognitive realism in ontogeny	31

2.3.7	D7: Explanation	32
2.4	Core developmental phenomena	34
2.4.1	The abstractness of early representations	34
2.4.2	Argument omission in early production	41
2.4.3	Argument-structure overgeneralization in early production	47
2.4.4	Explananda for a usage-based model of language acquisition	52
2.5	Computational usage-based models of language acquisition	53
2.5.1	Semantic-grammar models	53
2.5.2	Usage-based distributional models	60
2.5.3	A comparison	62
3	The Syntagmatic-Paradigmatic Learner	69
3.1	Introduction	69
3.2	General properties of input items to the model	70
3.2.1	Input items: utterances and conceptualizations of situations	70
3.2.2	The structure of the conceptual representations	71
3.3	Constructions	73
3.3.1	Constructions as representational primitives	73
3.3.2	A formal definition of constructions and the construction	74
3.4	Defining the space of possible analyses	77
3.4.1	Mapping constructions to situations	77
3.4.2	Three general constraints	80
3.4.3	Starting a derivation: concatenation	81
3.4.4	Ignoring words	82
3.4.5	Applying construction-mapping pairings	82
3.4.6	An example of the space of possible derivations	85
3.5	Selecting the best analysis	95
3.5.1	The probability model for derivations	95
3.5.2	Equivalent derivations: parses	98
3.5.3	An example of the probability model	100
3.5.4	Implementation: linear processing and pruning	103
3.5.5	SPL as a usage-based processing model	105
3.6	Learning	105
3.6.1	Reinforcement	106
3.6.2	Syntagmatization	109
3.6.3	Paradigmatization	111
3.6.4	Cross-situational learning	115
3.6.5	SPL as a usage-based learner	118
3.7	Generation	120
3.7.1	Differences with the analysis procedure	120
3.7.2	Expressivity	121
3.7.3	Selecting the best analysis and utterance	122

3.7.4	An example of the generation procedure	122
3.8	Meeting desiderata with SPL	124
4	Modeling the acquisition of meaning	129
4.1	Three problems in acquiring meaning	130
4.2	The informativeness of the situation	132
4.2.1	Earlier research	132
4.2.2	How available are the communicated concepts	137
4.2.3	Noise-reduction through understanding intentionality	144
4.2.4	Interpretation and implications	149
4.2.5	The issue of situational interdependence	150
4.2.6	Discussion	153
4.3	Towards a realistic simulation procedure	155
4.3.1	Earlier methods	155
4.3.2	Operationalization of the input generation procedure	157
4.4	Directions for modeling symbol acquisition	162
5	Comprehension experiments	165
5.1	Measuring comprehension	165
5.1.1	General evaluation	166
5.1.2	Evaluating the used representations	167
5.2	Global evaluation	167
5.2.1	Identification	167
5.2.2	Utterance coverage	168
5.2.3	Situation coverage	171
5.2.4	Robustness to uncertainty and noise	173
5.3	Used representations	174
5.3.1	The use of chunks	174
5.3.2	The use of bootstrapping	176
5.3.3	The use of concatenation	179
5.3.4	The length and abstraction of the used representations	180
5.4	Desiderata and explananda	185
6	Entering the black box	189
6.1	Learning mechanisms	189
6.1.1	Lexical learning	190
6.1.2	Grammatical learning	192
6.2	The representational potential	194
6.2.1	Length of the acquired constructions	194
6.2.2	Abstraction in the representational potential	199
6.3	The independence of morphemes	201
6.3.1	Entity words	203
6.3.2	Attribute words	204
6.3.3	Pronouns	206
6.3.4	Event words	206

6.3.5	Role-marking words	209
6.3.6	Comparing the classes	210
6.3.7	Discussion	211
6.4	The growth of the caused-motion construction	213
6.5	Discussion	216
7	Production experiments	221
7.1	Global development of production	221
7.1.1	Evaluation	221
7.1.2	Results	222
7.1.3	An example	224
7.1.4	Robustness to uncertainty and noise	225
7.2	Error analysis	227
7.2.1	Lexical errors	228
7.2.2	Argument structure errors	231
7.2.3	Argument omission	232
7.3	Overgeneralization	234
7.3.1	Motivation and Experimental set-up	234
7.3.2	Results	235
7.3.3	Factors in the overgeneralization and retreat	240
7.4	Discussion	241
8	Concluding remarks	243
8.1	Recapitulating SPL	244
8.2	The behavior of SPL	246
8.3	The representations acquired by SPL	248
8.4	Desiderata and explananda	249
8.5	Suggestions for the usage-based conception	252
8.6	Suggestions for cognitive modeling	253
	Bibliography	255
	Summary	269
	Samenvatting	277
	Curriculum Vitæ	285

Acknowledgements

Throughout the (almost) five years that I have been working on the project constituting this dissertation, I have had the pleasure of being surrounded by many great people who deserve to be acknowledged.

The first words of gratefulness definitely go out to my supervisors, Arie Verhagen and Rens Bod. With each of them packing a vast breadth and depth of knowledge, combined with their highly original and thorough ways of thinking, I could not have wished for a more nurturing home for my ideas.

Second, I would like to thank all other teachers and mentors that have fulfilled important exemplary roles in various phases of my academic development, both before and during my doctoral research. I would like to thank, in a somewhat chronological order, Gé Vaartjes, Leo van Santen, Ronny Boogaart, Cor van Bree, Ariane van Santen, Marijke van der Wal, Ton van der Wouden, Felix Ameka, Egbert Fortuin, Sandy Thompson, Pat Clancy, Jack DuBois, Stef Grondelaers, Remko Scha, Jelle Zuidema, Melissa Bowerman, Afsaneh Fazly, and Suzanne Stevenson for sharing their insights and stimulating my academic growth.

Important parts of this research have been discussed with Libby Barak, Gideon Borensztajn, Ailis Cournane, Max van Duijn, Stewart McCauley, Aida Nematzadeh, and Gareth O'Neill, whose criticisms helped further my thinking and writing and whose contributions I would like to acknowledge. The comments of the members of the committee, Afra Alishahi, Roberta D'Alessandro, and Ewa Dąbrowska, stimulated me to dot the i's and cross the t's, for which I would like to thank them.

I am very grateful for the academic environments I have spent time in over the past years. Many thanks go out to the participants of the CLC-lab and the LaCo-group at the ILLC, University of Amsterdam, colleagues at the LUCL, and the members of the SuzGroup and, wider, the CL-group, during my stay at the University of Toronto.

The research reported in chapter 4 could not have been carried out without the help of my two research assistants, Eva Rieborn and Marten van der

Meulen, who laboriously and conscientiously coded the caregiver-child interactions. Those caregiver-child interactions would not have been there at all if it weren't for Marinus van IJzendoorn and Marian Bakermans-Kranenburg of the department of Child Studies at Leiden University. Their generosity, in allowing me to use their data as well as occupying a desk in the department to digitize the material, is gratefully acknowledged. The input-generation procedure described in the same chapter is an extension of that of Afra Alishahi, who generously allowed me to use her code.

In the academic years 2013-2014 and 2014-2015, I lectured at the department of Dutch Language. I would like to thank my colleagues there, Ronny Boogaart, Gijsbert Rutten, Ariane van Santen, Tanja Simons, Arie Verhagen, and Marijke van der Wal for helping me out, as an inexperienced teacher, and providing me with an inspiring environment.

The research reported in this dissertation could not have been carried out without the generous *Promoties in de Geesteswetenschappen* grant of the NWO, whose support I would like to acknowledge.

On a personal level, there are too many to be thanked. I am grateful for my family, friends, and (former) partners for all the support, wisdom, joy, and love I have experienced. A warm thank you to all of you.

One person definitely deserves the 'last-but-not-least': my dear friend Folgert Karsdorp. During times of academic and personal misery his support has made the difference. Folgert's contributions at various stages of my dissertation research have been numerous and our conversations shaped chapters 5-7 to a large extent.

All remaining imperfection is solely my own.