



UvA-DARE (Digital Academic Repository)

Neural syntax

Fitz, H.

Publication date
2009

[Link to publication](#)

Citation for published version (APA):

Fitz, H. (2009). *Neural syntax*. [Thesis, fully internal, Universiteit van Amsterdam]. Institute for Logic, Language and Computation.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Contents

Acknowledgments	xi
1 Introduction	1
1.1 Computational models in language acquisition	1
1.2 Why neural networks?	4
1.3 Thesis outline	5
2 Recursion in neural networks	9
2.1 Introduction	9
2.2 The computational power of neural networks	10
2.2.1 Feed-forward networks	10
2.2.2 Recurrent networks	11
2.3 RNNs and string language learning	14
2.3.1 Regular languages	16
2.3.2 Context-free languages	16
2.3.3 Context-sensitive languages	18
Remark on backpropagation	20
2.4 RNNs and complex sentence structure	21
2.4.1 Recursive auto-associative memory	22
RAAM applications, extensions, and limitations	25
2.4.2 A subsymbolic parser for embedded clauses	26
The plausibility of SPEC	30
2.4.3 Simple recurrence and complex structure	31
2.4.4 Structural processing with semantic constraints	34
The effect of semantic constraints	36
2.4.5 Increasing grammatical complexity	37
Discussion of Christiansen's results	39
2.5 Summary	43

3	The Dual-path model	45
3.1	Limitations of the SRN approach	45
3.1.1	Meaning	45
3.1.2	Definiteness	48
3.1.3	Symbolic generalization	51
3.2	Features of the Dual-path model	52
3.2.1	Separate pathways	52
3.2.2	The WHAT-WHERE division	53
3.2.3	Event semantics	54
3.3	Dual-path model architecture explained	55
3.3.1	Sequencing system	55
3.3.2	Message-lexical system	57
3.3.3	Event semantics-layer	60
3.4	Production sample	62
3.5	Model assumptions	66
3.5.1	Learning assumptions	66
	Learning-as-processing	66
	Prediction error	67
3.5.2	Architectural assumptions	68
	Two pathways	68
	SRN sequencing	68
3.5.3	Representational assumptions	68
	WHAT-WHERE separation	68
	XYZ roles	69
3.6	Past research with the Dual-path model	70
4	Learning	73
4.1	Introduction	73
4.2	Artificial language and method	76
4.2.1	Artificial language	76
4.2.2	Method	79
4.3	Message representation comparison	80
4.3.1	Random baseline	80
4.3.2	Simple-event message	81
4.3.3	Event-order message	82
4.3.4	Event-link message	88
4.3.5	Event-order-link message	91
4.3.6	Binding message	94
4.3.7	Topic-focus message	96
4.3.8	Simple topic-focus message	98
4.3.9	Summary message types	100
4.4	The gapped element	104
4.4.1	Ambiguities	105

4.4.2	Acquisition	106
4.4.3	Alternations	108
4.4.4	Performance comparison <i>gap-link</i> versus <i>no gap-link</i>	109
4.5	Discussion	112
5	Model analysis	115
5.1	Behavioral analysis	115
5.1.1	Syntactic alternations	115
	Bigram statistics	119
	Frequency of semantic roles	120
	Error analysis	121
5.1.2	Topic and focus	123
5.2	Representational analysis	128
5.2.1	Lexical categories	129
5.2.2	Thematic role sequencing	132
	Hierarchical planning	136
5.2.3	Phrasal categories and argument structure	138
	Linear discriminant analysis	138
5.2.4	Verb, noun and prepositional phrases	138
5.2.5	Argument structure	143
5.2.6	Clause-level analysis	153
	Principal components analysis	153
	Clause type comparison	154
6	Generalization	161
6.1	Introduction	161
6.2	Knowledge transfer	161
6.3	Extension of language and semantics	163
6.4	Structural generalization	165
6.5	Lexical generalization	169
6.6	Right-branching versus center-embedding	174
6.7	Recursive productivity	184
6.7.1	Is recursion uniquely human?	184
6.7.2	Recursion and productivity	186
6.7.3	The innateness of recursion	187
6.7.4	A proper <i>explanandum</i>	188
6.7.5	Recursive generalization in the Dual-path model	190
6.8	Conclusion	200
7	Learning polar interrogatives	203
7.1	The general controversy	203
7.1.1	The learning problem	205
7.1.2	Three empirical hypotheses	207

7.1.3	Statistical learning	212
	The Lewis & Elman model	214
	The Reali & Christiansen model	216
7.2	The recursive Dual-path model approach	218
7.2.1	Artificial language, semantics and method	220
7.3	Modelling results	222
7.3.1	Simple polar interrogatives	223
7.3.2	Complex <i>wh</i> -questions	224
	Error analysis	226
7.3.3	Bootstrapping	230
7.4	Structure-dependence revisited	237
7.4.1	Hidden layer analysis	241
	Hierarchical structure	241
	Generalization by analogy	244
	Questions and declaratives	245
7.5	Discussion	247
8	The accessibility hierarchy	253
8.1	Introduction	253
8.2	The accessibility hierarchy in development	257
8.3	The recursive Dual-path model approach	258
8.4	Language and method	259
8.5	The accessibility hierarchy in the recursive Dual-path model	260
	8.5.1 The S>A contrast	261
	8.5.2 The A>P contrast	265
	8.5.3 P-, IO-, and OBL-relatives	267
8.6	Eliminating the relative clause hierarchy	268
8.7	Discussion	270
9	Conclusions	273
9.1	Key findings	273
	9.1.1 Learning as transduction	273
	9.1.2 Generalization with semantic similarities	274
	9.1.3 Differential processing due to input factors	275
9.2	Future directions	276
	9.2.1 Perspective taking	276
	9.2.2 Cross-linguistic study	278
	9.2.3 Dynamic message	280
A	Dual-path model details	283
A.1	Model specification	283
A.2	Training procedure	284
A.3	Simulation environment	285

B Improved question learning	287
B.1 New data	287
B.2 Analysis	287
List of Abbreviations	289
Bibliography	291
Samenvatting	313