

# **Mathematical Structure of Syntactic Merge**

An Algebraic Model for Generative Linguistics

Matilde Marcolli, Noam Chomsky, and Robert C. Berwick

The MIT Press  
Cambridge, Massachusetts  
London, England

# Contents

|  |           |
|--|-----------|
| Series Foreword  | xv        |
| List of Figures  | xvii      |
| Editor's Note  | xix       |
| <b>0 Minimalism and Merge: Introduction</b>            | <b>1</b>  |
| 0.1 Why This Monograph?                                | 2         |
| 0.2 Key Ideas  | 5         |
| 0.3 Summary of Main Linguistic Results                 | 11        |
| 0.4 Acknowledgment                                     | 16        |
| <b>1 The Mathematical Structure of Syntactic Merge</b> | <b>19</b> |
| 1.1 A Mathematical Model of Syntactic Merge            | 19        |
| 1.1.1 Notational Conventions                           | 19        |
| 1.1.2 Syntactic Objects and the Merge Magma            | 22        |
| 1.1.3 On the Use of the Tree Notation                  | 23        |
| 1.1.3.1 Planarity and lists versus sets                | 25        |
| 1.1.3.2 Algebra and operations on trees                | 26        |
| 1.2 Workspaces: Product and Coproduct                  | 27        |
| 1.2.1 Combinatorial Hopf Algebras and Workspaces       | 36        |
| 1.3 Action of Merge on Workspaces                      | 40        |
| 1.3.1 Matching Terms                                   | 40        |
| 1.3.2 Grafting   | 41        |
| 1.3.3 The Merge Operators                              | 42        |
| 1.3.4 Examples   | 46        |
| 1.4 Forms of Merge                                     | 48        |
| 1.4.1 Different Forms of Merge                         | 48        |
| 1.4.2 External Merge                                   | 49        |

|         |  |     |
|---------|--|-----|
| 1.4.3   | Internal Merge                                 | 49  |
| 1.4.3.1 | Is Internal Merge composite?                   | 50  |
| 1.4.4   | Internal Merge: An Example                     | 52  |
| 1.4.5   | Sideward Merge, Cases 2(b) and 3(b)            | 54  |
| 1.4.6   | Sideward Merge, Case 3(a)                      | 54  |
| 1.5     | Minimal Search                                 | 56  |
| 1.5.1   | The Minimality of Minimal Search               | 56  |
| 1.5.2   | Weighted Terms in the Coproduct                | 58  |
| 1.5.3   | Minimal Search: Internal and External Merge    | 60  |
| 1.6     | Other Linguistic Properties                    | 61  |
| 1.6.1   | Minimal Yield and Complexity                   | 62  |
| 1.6.2   | Cases of Merge and Size Counting               | 65  |
| 1.6.3   | No Complexity Loss Constraint                  | 72  |
| 1.7     | Counter-cyclicity and Extensions of Merge?     | 73  |
| 1.7.1   | Premise: Lie Algebras and Hopf Algebras        | 75  |
| 1.7.2   | Insertion Lie Algebra                          | 77  |
| 1.7.3   | Insertion Lie Algebra of Workspaces            | 78  |
| 1.8     | Cancellation of Copies                         | 81  |
| 1.9     | Merge Is Markovian                             | 86  |
| 1.10    | The Core Computational Structure of Merge      | 92  |
| 1.11    | Constraints on Merge: The $n$ -arity Question  | 95  |
| 1.11.1  | The $n$ -ary Merge Magma                       | 96  |
| 1.11.2  | Undergeneration                                | 98  |
| 1.11.3  | The Structure of a Hypothetical $n$ -ary Merge | 100 |
| 1.11.4  | Overgeneration                                 | 102 |
| 1.12    | A Model of Externalization                     | 105 |
| 1.12.1  | Externalization: A Preliminary Discussion      | 105 |
| 1.12.2  | Externalization as Correspondence              | 108 |
| 1.12.3  | Correspondences                                | 111 |
| 1.12.4  | Magma and Nonassociative Algebra               | 113 |
| 1.12.5  | Merge Representation                           | 114 |
| 1.12.6  | Externalization Correspondence and Algebras    | 117 |
| 1.12.7  | The Role of Syntactic Parameters               | 119 |
| 1.13    | Externalization and Planarization              | 122 |
| 1.13.1  | Commutative and Noncommutative Magmas          | 123 |
| 1.13.2  | Planarization versus Externalization           | 125 |
| 1.13.3  | Abstract Head Functions                        | 128 |

|          |  |            |
|----------|--|------------|
| 1.14     | Phase Theory   | 131        |
| 1.15     | Labeling Algorithm   | 139        |
| 1.16     | Form Set   | 141        |
| 1.16.1   | FormSet and Workspaces                                       | 145        |
| 1.16.2   | FormSet: Not an $n$ -ary Merge                               | 146        |
| 1.17     | Merge and Fundamental Combinatorial Recursions<br>in Physics | 147        |
| 1.17.1   | The Recursive Construction of Dyson–Schwinger<br>Equations   | 149        |
| <b>2</b> | <b>Minimalism Old and New: A Hopf Algebra Comparison</b>     | <b>153</b> |
| 2.1      | Introduction   | 153        |
| 2.2      | Old Minimalism and Loday–Ronco Hopf Algebra                  | 156        |
| 2.2.1    | Preliminaries on Hopf Algebras of Rooted<br>Trees            | 158        |
| 2.2.2    | Binary Rooted Trees and Admissible Cuts                      | 159        |
| 2.2.3    | The Loday–Ronco Hopf Algebra of Binary<br>Rooted Trees       | 159        |
| 2.2.4    | Graphical Form of Coproduct and Product                      | 161        |
| 2.2.5    | Hopf Algebra Comparisons                                     | 163        |
| 2.2.6    | Labeled Trees  | 165        |
| 2.2.7    | Old Formulation of External Merge and Internal<br>Merge      | 166        |
| 2.2.8    | Old Form of External Merge                                   | 168        |
| 2.2.9    | Old Version of Internal Merge                                | 168        |
| 2.2.10   | Old Version of Internal Merge and the Coalgebra<br>Structure | 170        |
| 2.2.11   | Old Version of Internal Merge and Algebra<br>Structure       | 172        |
| 2.2.12   | Iterated Internal Merge                                      | 173        |
| 2.2.13   | Feature Checking Complexity                                  | 175        |
| 2.2.14   | Coideals, Recursion, and Symmetries                          | 178        |
| 2.2.15   | Partial Operated Algebra                                     | 179        |
| 2.2.16   | Old Version of External Merge and Hopf Algebra<br>Structure  | 180        |
| 2.3      | Summary Comparison of Old Minimalism and New<br>Minimalism   | 182        |
| 2.3.1    | The Core Computational Structure                             | 183        |

|          |   |            |
|----------|---|------------|
| 2.3.1.1  | Core process of New Minimalism                                | 183        |
| 2.3.1.2  | Core process of Old Minimalism                                | 183        |
| 2.3.2    | Relation to Combinatorial Dyson–Schwinger Equations           | 184        |
| 2.3.2.1  | DS equations and New Minimalism                               | 184        |
| 2.3.2.2  | DS equations and Old Minimalism                               | 185        |
| 2.3.3    | Combinatorial Objects   | 186        |
| 2.3.3.1  | Syntactic objects and workspaces in New Minimalism            | 186        |
| 2.3.3.2  | Labeled planar trees  | 188        |
| 2.3.4    | Action of Merge   | 188        |
| 2.3.4.1  | Action of Merge in New Minimalism                             | 189        |
| 2.3.4.2  | Action of Merge in Old Minimalism                             | 191        |
| 2.3.4.3  | Labeling in Old Minimalism and New Minimalism                 | 192        |
| 2.4      | Conclusions   | 194        |
| 2.5      | Comparison with Other Models                                  | 196        |
| 2.5.1    | Tree-Adjoining Grammars                                       | 196        |
| 2.5.2    | Tensor Models   | 198        |
| 2.5.3    | Physics Methods in Minimalism                                 | 199        |
| <b>3</b> | <b>The Syntax–Semantics Interface: An Algebraic Model</b>     | <b>201</b> |
| 3.1      | Modeling the Syntax–Semantics Interface                       | 201        |
| 3.1.1    | Some Conceptual Requirements for a Syntax–Semantics Interface | 203        |
| 3.1.2    | Syntax  | 204        |
| 3.1.2.1  | Remark on the Hopf Algebra coproduct                          | 206        |
| 3.1.3    | Consistency and Substructures: Preliminary Discussion         | 207        |
| 3.1.4    | Algebraic Renormalization: A Short Summary                    | 208        |
| 3.1.5    | Semantic Spaces   | 214        |
| 3.1.5.1  | Neuroscience data and syntax–semantics interface models       | 215        |
| 3.1.5.2  | Formal properties of semantic spaces                          | 216        |
| 3.1.5.3  | Concept spaces outside of language                            | 217        |
| 3.2      | Syntax–Semantics Interface as Renormalization: Toy Models     | 218        |

|         |  |     |
|---------|--|-----|
| 3.2.1   | A Simple Toy Model: Head-Driven Syntax–Semantics Interface       | 219 |
| 3.2.2   | Head-Driven Interfaces and Convexity                             | 226 |
| 3.2.2.1 | Comparison functions   | 227 |
| 3.2.2.2 | Threshold Rota-Baxter operators                                  | 228 |
| 3.2.2.3 | Viterbi-valued semiring character                                | 229 |
| 3.2.2.4 | Birkhoff factorization with threshold operators                  | 231 |
| 3.2.2.5 | Convex neighborhoods   | 232 |
| 3.2.3   | Head-Driven Interfaces and Vector Models                         | 234 |
| 3.2.3.1 | Max-plus-valued semiring character                               | 234 |
| 3.2.3.2 | Hyperplane arrangements  | 235 |
| 3.2.3.3 | ReLU Birkhoff factorization                                      | 236 |
| 3.2.4   | Not a Tensor-Product Semantics                                   | 237 |
| 3.2.5   | Boolean Semiring   | 238 |
| 3.3     | The Image of Syntax Inside Semantics                             | 240 |
| 3.4     | Head Functions, Moduli Spaces, Associahedra, and Externalization | 243 |
| 3.4.1   | Preliminary Discussion   | 243 |
| 3.4.2   | Associahedra and Moduli Spaces                                   | 246 |
| 3.4.3   | Head Functions, Convex Semantic Spaces, and Metric Trees         | 251 |
| 3.4.4   | Origami Folding and Externalization                              | 253 |
| 3.4.5   | An Example   | 255 |
| 3.4.6   | Geometric View of Planarization Questions                        | 260 |
| 3.4.6.1 | Kayne’s LCA algorithm  | 260 |
| 3.4.6.2 | Cinque’s Abstract Functional Lexicon                             | 260 |
| 3.4.6.3 | Geometric view of syntactic parameters                           | 261 |
| 3.5     | Birkhoff Factorization and (Semi)ring Parsing                    | 262 |
| 3.5.1   | Preliminary Discussion   | 263 |
| 3.5.2   | Minimal Yield as Birkhoff Factorization                          | 266 |
| 3.5.2.1 | Effect of Merge on workspaces                                    | 266 |
| 3.5.2.2 | Laurent series ring of Merge derivations                         | 268 |
| 3.5.3   | Birkhoff Factorization in Algebroids                             | 273 |
| 3.5.4   | Bialgebroids and Rota-Baxter Algebroids                          | 276 |
| 3.5.4.1 | Algebroids and directed graph schemes                            | 276 |
| 3.5.4.2 | Rota–Baxter algebroids   | 278 |

|          |  |            |
|----------|--|------------|
| 3.5.4.3  | Rota–Baxter semiringoids                                       | 280        |
| 3.5.4.4  | Birkhoff factorization in algebroids and semiringoids          | 281        |
| 3.5.5    | Parsing Semirings and Merge Derivations                        | 284        |
| 3.5.6    | Summary of Semiringoid Parsing                                 | 285        |
| 3.6      | Pietroski’s Compositional Semantics                            | 287        |
| 3.6.1    | Combine Operation in Pietroski’s Semantics                     | 289        |
| 3.6.2    | Concatenate Operation  | 290        |
| 3.6.2.1  | The role of idempotents  | 296        |
| 3.6.2.2  | An example   | 297        |
| 3.6.3    | Predicate Saturation and Operadic Structure                    | 299        |
| 3.6.3.1  | Syntactic objects and operads                                  | 299        |
| 3.6.3.2  | Semantic spaces and operads                                    | 301        |
| 3.6.3.3  | Syntax-driven compositional semantics                          | 302        |
| 3.7      | Adjunction, Embedded Constructions, and the Pair-Merge Problem | 303        |
| 3.7.1    | Adjunctions and Hierarchical Structures                        | 306        |
| 3.8      | Language-Specific Conditions and Theta Theory                  | 313        |
| 3.8.1    | Theta Theory   | 314        |
| 3.8.2    | Deriving Obligatory Control                                    | 320        |
| 3.9      | Heim–Kratzer Semantics   | 326        |
| 3.9.1    | Boolean Characters and Interpretability                        | 328        |
| 3.9.2    | Topologizing Semantic Types                                    | 330        |
| 3.9.3    | Boolean Probes and Topological Types                           | 334        |
| 3.9.4    | Fuzzy Semantic Types   | 335        |
| 3.10     | No, They Don’t: Transformers as Characters                     | 336        |
| 3.10.1   | Attention Modules  | 338        |
| 3.10.1.1 | Heads and heads  | 340        |
| 3.10.2   | Maximizing Attention   | 340        |
| 3.10.3   | Attention-Detectable Syntactic Relations                       | 341        |
| 3.10.4   | Rota-Baxter Structures and Attention                           | 342        |
| 3.10.5   | Syntax as an Inverse Problem: Physics as Metaphor              | 344        |
| <b>4</b> | <b>Summary of Mathematical Concepts</b>                        | <b>347</b> |
| 4.1      | Categories   | 347        |
| 4.2      | Hopf Algebras  | 349        |

|       |   |     |
|-------|---|-----|
| 4.2.1 | Main Idea                                   | 349 |
| 4.2.2 | Formal Definition                           | 351 |
| 4.2.3 | Bialgebras and Hopf Algebras                | 353 |
| 4.3   | Rooted Trees                                | 354 |
| 4.4   | Other Algebraic Structures                  | 355 |
| 4.4.1 | Main Idea                                   | 355 |
| 4.4.2 | Rings, Algebras, Ideals, Modules, Semirings | 356 |
| 4.5   | Point Set Topology                          | 358 |
| 4.6   | Further Remarks on Birkhoff Factorization   | 360 |
|       | Bibliography                                | 367 |
|       | Index                                       | 381 |