

# Table of contents

 <https://doi.org/10.1075/nlp.10.toc>

Pages v–viii of

**Modelling Language**

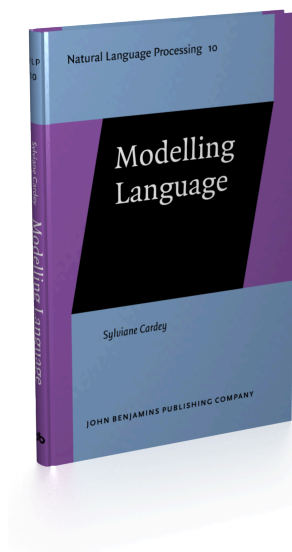
**Sylviane Cardey**

[**Natural Language Processing**, 10] 2013. xi, 194 pp.

© John Benjamins Publishing Company

This electronic file may not be altered in any way. For any reuse of this material written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

For further information, please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website at [benjamins.com/rights](http://benjamins.com/rights)



# Table of contents

|                                                    |    |
|----------------------------------------------------|----|
| Preface                                            | IX |
| Prologue                                           | XI |
| Introduction                                       | 1  |
| <b>PART 1. System, language and its components</b> |    |
| CHAPTER 1.1                                        |    |
| <b>The concept of system</b>                       | 5  |
| 1.1.1 System                                       | 5  |
| 1.1.2 Systemicity                                  | 5  |
| CHAPTER 1.2                                        |    |
| <b>Language as a system</b>                        | 9  |
| 1.2.1 Grammatical system                           | 10 |
| 1.2.2 Language typology                            | 11 |
| 1.2.3 Lexicology, morphology and syntax            | 13 |
| CHAPTER 1.3                                        |    |
| <b>The system's micro-components</b>               | 17 |
| 1.3.1 The word                                     | 17 |
| 1.3.2 Morphemes and syllables                      | 20 |
| 1.3.3 Parts of speech                              | 21 |
| CHAPTER 1.4                                        |    |
| <b>Syntactic analysis</b>                          | 25 |
| CHAPTER 1.5                                        |    |
| <b>Semantics</b>                                   | 27 |
| CHAPTER 1.6                                        |    |
| <b>Norm in language</b>                            | 29 |
| 1.6.1 Synchrony and diachrony                      | 29 |
| 1.6.2 Good usage                                   | 31 |

**PART 2. Modelling the norms****CHAPTER 2.1****Model 35****CHAPTER 2.2****Our model 37**

- 2.2.1 The linguistic model 37
  - 2.2.1.1 Macrocosmic representation 37
  - 2.2.1.2 Microscopic approach to morphology 41
    - 2.2.1.2.1 Nested elements 43
      - 2.2.1.2.1.1 Morpho-syntax 45
        - 2.2.1.2.1.2 Lexical morphology 48
          - 2.2.1.2.1.2.1 Derivation 48
          - 2.2.1.2.1.2.2 Composition 49
  - 2.2.1.3 Systemic linguistic modelling of other languages 51
    - 2.2.1.3.1 Malay 51
    - 2.2.1.3.2 Arabic 52
  - 2.2.1.4 Concept of micro-system 53
    - 2.2.1.4.1 Algorithmic micro-system 53
    - 2.2.1.4.2 Examples of micro-systems 55
      - 2.2.1.4.2.1 Algorithmic micro-system example 1 and its various representations. French words starting with 'ap' 55
        - 2.2.1.4.2.1.1 The algorithm 56
        - 2.2.1.4.2.1.2 Representations 58
      - 2.2.1.4.2.2 Algorithmic micro-system example 2. English words ending with -ed, -ing, -er, -est, -en 61
      - 2.2.1.4.2.3 Algorithmic micro-system example 3. The agreement of the French past participle 63
      - 2.2.1.4.2.4 Algorithmic micro-system example 4. Declension of German adjectives 65
  - 2.2.1.5 Our model for syntax 71
    - 2.2.1.5.1 Syntax and morphology 72
    - 2.2.1.5.2 Lexical syntax 76
  - 2.2.1.6 The same formal representation over domains and languages 76
    - 2.2.1.6.1 Interlanguage norms 78
    - 2.2.1.6.2 Divergent structures 79

|             |                                                                 |     |
|-------------|-----------------------------------------------------------------|-----|
| 2.2.1.7     | Disambiguation                                                  | 83  |
| 2.2.1.7.1   | Disambiguation in English                                       | 83  |
| 2.2.1.7.2   | How to represent a problem of ambiguity<br>in composition: Thai | 83  |
| 2.2.1.7.3   | Disambiguating Chinese and Thai                                 | 85  |
| 2.2.2       | The mathematical model                                          | 87  |
| 2.2.2.1     | Necessary notions                                               | 87  |
| 2.2.2.1.1   | Set                                                             | 88  |
| 2.2.2.1.2   | Partition                                                       | 91  |
| 2.2.2.1.3   | Logical operations                                              | 93  |
| 2.2.2.1.4   | Binary relations                                                | 93  |
| 2.2.2.1.5   | Modelling algorithmic operations                                | 96  |
| 2.2.2.2     | Mathematical modelling of micro-systemic linguistics            | 96  |
| 2.2.2.2.1   | Establishing a systemic linguistic analysis                     | 97  |
| 2.2.2.2.2   | Establishing the partitions                                     | 97  |
| 2.2.2.2.2.1 | The non-contextual analysis                                     | 98  |
| 2.2.2.2.2.2 | The in-context analysis                                         | 102 |
| 2.2.2.2.3   | Set analysis                                                    | 104 |
| 2.2.2.2.3.1 | Proper subsetting                                               | 105 |
| 2.2.2.2.3.2 | Disjunction                                                     | 105 |
| 2.2.2.2.4   | In-context analysis reviewed                                    | 106 |
| 2.2.2.2.5   | Formulation of the super-system                                 | 106 |
| 2.2.2.3     | Optimisation considerations                                     | 108 |
| 2.2.2.4     | Applying the abstract mathematical model                        | 109 |
| 2.2.2.4.1   | Model-driven evaluation                                         | 110 |

## PART 3. Methodologies and applications

### CHAPTER 3.1

|                  |     |
|------------------|-----|
| Grammar checkers | 115 |
|------------------|-----|

### CHAPTER 3.2

|                       |     |
|-----------------------|-----|
| Part of speech tagger | 121 |
|-----------------------|-----|

|         |                                     |     |
|---------|-------------------------------------|-----|
| 3.2.1   | Morphological rule dictionary       | 121 |
| 3.2.2   | Labelgram                           | 127 |
| 3.2.3   | Applications to different languages | 128 |
| 3.2.3.1 | German                              | 128 |
| 3.2.3.2 | Spanish                             | 130 |
| 3.2.3.3 | English                             | 131 |
| 3.2.3.4 | French                              | 131 |

|             |                                                                                       |     |
|-------------|---------------------------------------------------------------------------------------|-----|
| 3.2.4       | Benchmarking                                                                          | 133 |
| 3.2.5       | Neologisms and Jabberwocky                                                            | 134 |
| CHAPTER 3.3 |                                                                                       |     |
|             | <b>Sense mining</b>                                                                   | 137 |
| 3.3.1       | Classification                                                                        | 137 |
| 3.3.2       | Semigram                                                                              | 139 |
| 3.3.3       | Testing and the classification rate                                                   | 141 |
| 3.3.4       | Results over different languages                                                      | 142 |
| CHAPTER 3.4 |                                                                                       |     |
|             | <b>Controlled languages</b>                                                           | 145 |
| CHAPTER 3.5 |                                                                                       |     |
|             | <b>Intralanguage ambiguity</b>                                                        | 149 |
| CHAPTER 3.6 |                                                                                       |     |
|             | <b>MultiCoDiCT</b>                                                                    | 151 |
| CHAPTER 3.7 |                                                                                       |     |
|             | <b>Controlled language and machine translation</b>                                    | 155 |
| 3.7.1       | Divergent structures                                                                  | 156 |
| 3.7.2       | Lexical divergences                                                                   | 158 |
| 3.7.3       | Translation architecture                                                              | 159 |
| 3.7.4       | Including a new language pair: Russian to Chinese                                     | 165 |
| 3.7.5       | Tests                                                                                 | 166 |
| 3.7.6       | Tracing                                                                               | 166 |
| CHAPTER 3.8 |                                                                                       |     |
|             | <b>Oral</b>                                                                           | 171 |
| 3.8.1       | Controlling the oral                                                                  | 171 |
| 3.8.1.1     | Quasi-homophones and recognition                                                      | 172 |
| 3.8.1.2     | Generation and interpretation of language sounds                                      | 173 |
| 3.8.1.3     | Examples of 12 languages with problems at the level of phonemes compared with English | 174 |
| 3.8.1.4     | Other problems for future work                                                        | 179 |
|             | <b>Conclusion</b>                                                                     | 181 |
|             | <b>Epilogue</b>                                                                       | 183 |
|             | <b>References</b>                                                                     | 185 |
|             | <b>Index</b>                                                                          | 191 |