

University of Groningen

## Phonological grammar and frequency

Sloos, Marjoleine

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*  
2013

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Sloos, M. (2013). *Phonological grammar and frequency: an integrated approach*. s.n.

**Copyright**

Other than for strictly personal use, 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), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

---

*Phonological Grammar and Frequency:  
an Integrated Approach*

*Evidence from German, Indonesian and Japanese*

---



rijksuniversiteit  
 groningen



DFG Deutsche  
Forschungsgemeinschaft

The research reported on in this thesis has been carried out under the auspices of the Hermann Paul School of Language Sciences in Freiburg (Germany), the Netherlands National Graduate School of Linguistics (Landelijke Onderzoekschool Taalwetenschap LOT), and the Center for Language and Cognition Groningen (CLCG) of the Faculty of Arts of the University of Groningen. This research was financially supported by the Deutsche Forschungsgemeinschaft (DFG) and the University of Groningen. The publication of this dissertation was financially supported by the University of Groningen.



Groningen Dissertations in Linguistics 111

ISBN: 978-90-367-6066-9

ISSN: 0928-0030

© 2013 by Marjoleine Sloos, all rights reserved.

Cover illustration: 'Gradience and categoricity'  
Acrylic and oil on canvas 60x60cm  
Christiaan Sloos  
© 2013.

Printed by CPI Koninklijke Wöhrmann, Zutphen, The Netherlands.

RIJKSUNIVERSITEIT GRONINGEN

# **Phonological Grammar and Frequency: an Integrated Approach**

**Evidence from German, Indonesian and  
Japanese**

**Proefschrift**

ter verkrijging van het doctoraat in de  
letteren  
aan de Rijksuniversiteit Groningen  
op gezag van de  
Rector Magnificus, dr. E. Sterken,  
in het openbaar te verdedigen op  
donderdag 28 februari 2013  
om 16.15 uur

door

Marjoleine Sloos  
geboren op 5 januari 1970  
te Woerden

Promotores: Prof. dr. ir. J. Nerbonne  
Prof. dr. J.M. van de Weijer

Copromotores: Dr. D.G. Gilbers  
Dr. C.S. Gooskens

Beoordelingscommissie: Prof. dr. V.J.J.P. van Heuven  
Prof. dr. M. van Oostendorp  
Prof. dr. G. Seiler

ISBN: 978-90-367-5989-2

*Voor mijn vader*







# Contents

Acknowledgements.....	xiii
List of symbols, glosses, and abbreviations.....	xvii
List of Figures.....	xvix
List of Tables .....	xxi

## Part I: Introduction & Methodology

<b>1 Towards hybrid modelling: Optimality Theory &amp; Exemplar Theory .....</b>	<b>3</b>
1.1 From grammar to usage.....	8
1.1.1 <i>Optimality Theory</i> .....	8
1.1.2 <i>Typology</i> .....	10
1.1.3 <i>The Emergence of the Unmarked</i> .....	11
1.1.4 <i>Frequency effects in Optimality Theory</i> .....	13
1.2 From usage to grammar .....	15
1.2.1 <i>Exemplar Theory</i> .....	15
1.2.2 <i>Usage-based grammar</i> .....	18
1.3 Towards hybrid modelling .....	19
1.3.1 <i>Prototypes</i> .....	20
1.3.2 <i>EPOT (Exemplar-Prototype-Optimality Theory)</i> .....	22
1.3.3 <i>Alternative: Grammar as Selection (van de Weijer 2012)</i> .....	23
1.4 Research question and hypotheses.....	25
1.5 Structure of this thesis .....	26
<i>Märchen (Rückert)</i> .....	29

<b>2 Merger and Reversal: The Standard German long vowels &lt;ä&gt; and &lt;e&gt;</b> .....	<b>31</b>
2.1 Ghost phoneme or real phoneme? Standard German long vowel <ä> .....	33
2.1.1 Phoneme or no phoneme? .....	33
2.1.2 Long vowel <ä> and low <r> .....	37
2.1.3 Standard, nationalect, regiolect, dialect .....	38
2.2 The historical background and the merger of the long vowels <ä> and <e> .....	40
2.2.1 Umlaut .....	41
2.2.2 Other lines of development of the long vowel <ä> .....	42
2.2.3 The merger of the long vowel <ä> and long <e> .....	42
2.3 Variation and the role of stylistic variation, orthography, and dialect contact .....	45
2.3.1 Dialect contact .....	45
2.3.2 Stylistic variation .....	45
2.3.3 Orthography .....	47
2.4 The reversal of the merger of the long vowel <ä> and long <e> .....	47
2.4.1 The data .....	47
2.4.2 Measuring long vowel <ä> and long <e> .....	48
2.4.3 Results .....	50
2.5 Discussion .....	52
2.6 Conclusion .....	55
<b>3 Methodological considerations: avoiding the halo-effect</b> .....	<b>57</b>
3.1 Swiss Standard German and the long vowel <ä> .....	60
3.1.1 Swiss Standard German .....	60
3.1.2 SSG long vowel <ä> .....	61
3.2 Categorical analysis I .....	62
3.2.1 The data .....	62
3.2.2 The coders and intercoder (dis-)agreement .....	64
3.2.3 Results .....	64
3.3 The acoustic analysis .....	68
3.4 The degree of Swiss Standard German accent: Online elicitation .....	70
3.4.1 Approach .....	70

3.4.2 Results.....	71
3.5 The halo effect.....	73
3.5.1 Categorical analysis II, SSG accent included .....	73
3.5.2 The halo effect in detail.....	76
3.6 Discussion and conclusion .....	80

## Part II: Type III frequency effects

<b>4 Rendaku love and hatred: opaque morphological structure.....</b>	<b>85</b>
4.1 Rendaku .....	86
4.2 Lexical variation.....	88
4.3 Methodology .....	93
4.3.1. Selection of the roots.....	93
4.3.2 Selection of the compounds.....	95
4.3.3 Frequency.....	95
4.4 Results.....	96
4.5 Rendaku analysis in EPOT .....	98
4.6 Conclusion:.....	103
<b>5 The Alemannic variety of Standard German long &lt;ä&gt;: opaque lexical structure .....</b>	<b>105</b>
5.1 The Alemannic and Swabian long vowel <ä>.....	107
5.2 Approach.....	109
5.2.1 Shadowing and phoneme restoration amalgamation.....	111
5.2.2 Selection of the stimuli .....	111
5.2.3 Preparation of the stimulus material .....	115
5.2.4 Estimation the level of standardness.....	116
5.3 Chasing shadows: an experimental approach to variation .....	116
5.3.1 Subjects.....	116
5.3.2 Stimulus material .....	117
5.3.3 Procedure .....	117
5.4 Results .....	118
5.5 Discussion .....	122
5.6 Conclusion .....	124

## Part III: Changing lexicons, changing grammars

<b>6 Dutch loanwords in Indonesian: Frequency effects are sensitive to the grammar</b> .....	<b>127</b>
6.1 Loanword phonology of Indonesian .....	132
6.2 Methodology.....	136
6.2.1 Database.....	136
6.2.2 Frequency.....	136
6.2.3 Hypotheses .....	138
6.3 Results: Frequency effects are sensitive to the grammar .....	139
6.4 The initial and the final stage of loanword integration .....	142
6.4.1 Lexical OT.....	142
6.4.2 Coalescence in loanwords and native words: An OT analysis.....	144
6.5. Gradual change from loanword to nativized word .....	148
6.5.1 Loanwords and Type III frequency effects .....	149
6.5.2 Frequency effects are sensitive to the grammar .....	150
6.6 Discussion and conclusion.....	151
<b>7 German: One rule, two frequency effects</b> .....	<b>153</b>
7.1 Geographical distribution .....	155
7.2 The interaction of frequency and grammar.....	162
7.3 Usage-based grammatical change.....	167
7.3.1 Frequency effects and the changing prototype of <ä> .....	167
7.3.2 Grammatical stages.....	168
7.4 A restriction on GEN .....	172
7.5 Conclusion .....	174

## Part IV: Conclusion

<b>8 Discussion and conclusions</b> .....	<b>179</b>
8.1 Results.....	180
8.2 Exemplar-Prototype-Optimality-Theory (EPOT) .....	184
8.3 Conclusion .....	185

<b>Summary</b> .....	<b>187</b>
<b>Samenvatting</b> .....	<b>191</b>
<b>Zusammenfassung</b> .....	<b>195</b>
<b>Appendices</b> .....	<b>199</b>
Appendix A: The degree of accentedness of all Swiss speakers.....	199
Appendix B: Stimulus material experiment .....	200
Appendix C: Database Indonesian.....	202
Appendix D: Results frequency-grammar interactions in Standard German long <ä> ...	208
<b>References</b> .....	<b>211</b>
<b>Curriculum vitae</b> .....	<b>229</b>



## *Acknowledgements*

---

It goes without saying that the most important person in the life of a PhD student is his/her supervisor. Nobody could even wish him/herself a better supervisor than I had: Jeroen van de Weijer. Dear Jeroen, at this place, I would like to express my feelings of gratitude, appreciation, and respect for you, but grammar as well as lexicon miserably fail to do so in all depth. Since you have repeatedly proven to be able to read between the lines, I trust that, despite the limitations of language, you can guess the depth of what I am trying to convey.

Although the physical distance between you and me was large, your prompt answers to any of my questions, requests for review of my work, and problems that I wished to discuss, continuously gave me the impression that you were always present. Questions there were many, often several per day, but you never got tired of answering them. Instead, you enkindled my enthusiasm, dismissed my worries (even when not expressed), or started an inspiring discussion concerning linguistics, which I deeply appreciated. I often asked you to review my work, and your endless and extremely adequate comments, corrections, and ‘nasty questions’, were (and still are) sometimes hard to bear, but always encouraged me to search for improvement. During the most difficult time of my life, you were there to support me to an incredible extent, solving all practical problems that could reasonably (and even more than that) be solved. And when it was really impossible to find a solution, you still found a way to encourage me with your everlasting sense of humour. I feel very obliged that, with wisdom, righteousness, and practical advice, you dismissed my sorrows and distress such that I could concentrate on my dissertation. Moreover, you turned my problems into challenges that a PhD student really needs to move forwards. Further, I am also deeply grateful for your hospitality, your support, and the opportunities that you arranged for me during my stay in Shanghai.

A number of other people considerably contributed to my academic development and the improvement of my articles and this dissertation, for which I am very grateful: John Nerbonne, Guido Seiler, Vincent van Heuven, Dicky Gilbers, Charlotte Gooskens, and Adele Cheetham. John, my supervisor at Groningen University, always pressed me to finish this dissertation as soon as possible, filled serious gaps in my statistical knowledge, and elevated the level of my articles significantly. Guido Seiler, my supervisor in Freiburg, was really concerned about the circumstances I had to deal with, urged me to investigate the data in more detail, and made me more aware of the historical background of my research. Vincent van Heuven, my professor of phonetics and statistics at Leiden University, from the very first proposal towards the final manuscript, remained interested in my research, was always willing to discuss my work, and removed many doubts about the methodology. Dicky Gilbers and

Charlotte Gooskens, my co-supervisors at Groningen University, supported me by carefully reading my articles and by useful discussions. Adele Cheetham, my tutor in academic writing of the University of Salford, with encouraging feed-back and much red ink, truly helped me to improve my writing skills, and increased my pleasure in writing.

Without data, no linguistics: I am very much indebted to a large number of people who were involved in different parts of my research such as providing data, participation in experiments, data analysis, stimuli generation, and statistical support. First of all, I express my gratitude towards Stefan Kleiner of the *Institut für deutsche Sprache* in Mannheim, Germany, for making available to me part of the *Corpus Deutsch Heute* and sharing frequency and other information about the data. Furthermore, I want to thank all Alemannic participants in the experiment, whose names I cannot mention here, but of whom I have very special memories. I thank Friedel Scheer-Nahor of the *Mütersproochgsellschaft* ‘Alemannic Native Speakers Association’ in Freiburg for quickly providing the contact data of many of these speakers, as well as the anonymous respondents of the online surveys. I sincerely thank all colleagues who helped me with the annotations of the data: Raffaella Baechler, Katharina Baier, Esther Jiresch, Brenda Jung, Lea Nirk, Michael Schäfer, Mirjam Schuler, Caroline Schwarz, Florian Sieg, and Nadieh Wattel. Thanks to Jos Pacilly for the Praat-script which I used extensively and to Thilo Weber for the German translation of the summary. I am especially thankful to Beata Bieniek, my German teacher, who helped me through the most difficult first months in Freiburg by patiently explaining all kinds of language and communication problems, and who was also willing to be locked into the isolated booth in the scary basements of the university to make recordings for her crazy linguistics student. And also, I want to thank all people who helped me out when I again started panicking after having lost my way in R: Harald Baayen (for the introduction into the forests and the trees), Çağrı Çöltekin (for the graphics), Martin Hilpert (for explanations in English), Sascha Wolfer (for the most difficult beginnings), Christian Wolk (for the maps). This project has been partially supported by the DFG-grant 1624/1, which is gratefully acknowledged. Thanks also to the RefWorks support team of the University of Groningen for prompt replies in case of troubles.

I deeply enjoyed the joined projects which I participated in during these years, and I want to express my thanks to all co-authors: Aone van Engelenhoven for his sense of humour and optimism, Bernadette Kushartanthi for her patience and interest, Clemens Poppe for his critical attitude and perseverance (yes, I really liked it), Nicoline van der Sijs for her inspiration and enthusiasm.

A friend in need is a friend indeed. During the dark years of my PhD, I enjoyed laughter and humour with my friends Bahareh Soohani and Sander Sauv  in Leiden, who stimulated me to continue and helped me to put things in perspective. I enjoyed the company of my friends in Freiburg, who were willing to listen, to support, and share their linguistic thoughts with me: Raffaella Baechler, Verena Haser, Nikolay Khakimov, and Glenda Leung. Very special thanks go to my *chado*, Japanese tea ceremony, teacher, Ulrich Haas, the



prototype of a teacher who never loses his patience. His tea lessons forced me to forget my distress and be aware of the here and now only. I will always keep in my heart the remembrance of the times he treated me as his honoured guest, and those moments when the difference between teacher and student vanished in the sharing of a bowl of tea. Further, I feel deeply grateful to Boudewijn van den Berg, who introduced me to Groningen University, which eventually turned out to be a turning point in my PhD. Moreover, he highly motivated and inspired me by his swift thinking. I also want to thank Wenting 'Viola' Yu for her friendship in Shanghai and Leiden, my Chinese name, and the fact that she willingly spent a whole day on a careful check of this thesis at the very last moment (of course I remain responsible for any errors). I bear warm feelings to all of these true friends. Further, I am deeply thankful to the people who made me really feel at home, and who were there to share tea and conversation with: Opa und Oma Stoll (und Bennie) in Freiburg and Freda de Goed in Groningen.

Last but not least, I want to thank him who supported me throughout my whole life, who has always been open to anything new I started to learn and kept up with it, whether it was music, bird-watching, Japanese tea ceremony, or even.....linguistics. Grateful for his silent listening to my thoughts, grateful for his questions which forced me to better formulate my thoughts, grateful for his understanding, I dedicate this book to my father.



## *List of symbols, glosses, and abbreviations*

---

<...>	orthographic representation
...	lexical form (of an exemplar)
/.../	underlying form
[...]	phonetic form/surface form
σ	syllable
2/3SG	2 <sup>nd</sup> and 3 <sup>rd</sup> person singular
ACT.	Active
ANIM.	Animate
CAUS.	Causative
DIM.	Diminutive
PAST.	Past tense
PLUR.	Plural
SUBJ.	Subjunctive
ATR	Advanced Tongue Root
C <sub>1</sub>	Coder 1 (in chapter 3)
C <sub>2</sub>	Coder 2 (in chapter 3)
EPOT	Exemplar-Prototype-Optimality-Theory
ET	Exemplar Theory
F <sub>1</sub>	First formant
F <sub>2</sub>	Second formant
GS	Grammar as Selection (van de Weijer 2012)
HF	High-frequency
HG	Harmonic Grammar
IDS	Institut für deutsche Sprache (Institute for German language)
LF	Low-frequency
MF	Moderately frequent
MHG	Middle High German
NHG	New High German
NSG	Northern varieties of Standard German (as opposed to SSG, see below)
OHG	Old High German
OT	Optimality Theory
SSG	Swiss Standard German
TETU	The Emergence of The Unmarked



## List of Figures

---

- Figure 1.1. Analogical network connections in Exemplar Theory (Bybee 2002).
- Figure 2.1. Main dialect groups in German.
- Figure 2.2. Vowel diagram of the International Phonetic Alphabet.
- Figure 3.1. Sites in Switzerland, Bludenz and Vaduz, investigated for the pronunciation of the long vowel <ä>.
- Figure 3.2. Regression tree of the codings of Rater 1.
- Figure 3.3. Regression tree of the codings of Rater 2.
- Figure 4.1a. Exemplars and connections of the LF root *siro* and compounds with *siro*.
- Figure 4.1b. Exemplars and connections of the HF root *hue* and compounds with *hue*.
- Figure 5.1. Tree diagram of the pronunciation of long vowel <ä> in the Alemannic variety of Standard German: education and dialect.
- Figure 5.2. Tree diagram of the pronunciation of long vowel <ä> in the Alemannic variety of Standard German: frequency and umlaut.
- Figure 6.1. The most prominent lexical strata of Indonesian: Malay, Sanskrit, Arabic, Dutch, and English.
- Figure 6.2. Percentage of coalescence in Dutch loanwords in Indonesia.
- Figure 7.1a. The vowel height of long <ä>, spoken by older speakers in non pre-r context.
- Figure 7.1b. The vowel height of long <ä>, spoken by older speakers in pre-r context.
- Figure 7.1c. The vowel height of long <ä>, spoken by younger speakers in non pre-r context.
- Figure 7.1d. The vowel height of long <ä>, spoken by younger speakers in pre-r context.
- Figure 7.2. Tree diagram of the pronunciation of long vowel <ä> for East Middle German.
- Figure 7.3. Tree diagram of the pronunciation of long vowel <ä> for Low Saxon.



## *List of Tables*

---

- Table 1.1. Schema of the case studies in this dissertation.
- Table 2.1. Frequencies of the formants of German /ɛ:/ and /e:/ in Hz in Standard German (Valackai 1998).
- Table 2.2. Mean formant values of German /e:/ and /ɛ:/ and their percentual differences (based on Valackai 1998).
- Table 2.3. German vowel specification (Wiese 2000: 20).
- Table 2.4. Five different realizations of long <ä> in different stylistic levels in percentages.
- Table 2.5. Vowel height in Austro-Bavarian.
- Table 2.6. Vowel height in Swabian-Alemannic.
- Table 2.7. Vowel height in in High-Franconian.
- Table 3.1. The number of [e:] and [ɛ:] codings by Coder 1 and Coder 2.
- Table 3.2. Number and percentages of [e:] and [ɛ:] codings of Coder 1 (specified).
- Table 3.3. Categorical analysis I of Coder 1.
- Table 3.4. Number and percentages of [e:] and [ɛ:] codings of Coder 2 (specified).
- Table 3.5. Categorical analysis I of Coder 2.
- Table 3.6. The acoustic analysis.
- Table 3.7. Summary of the results of the acoustic and categorical analyses I.
- Table 3.8. The number of respondents of the online survey.
- Table 3.9. Results of the optimal mixed model predicting the degree of Swiss Standard German accent.
- Table 3.10. Categorical analysis II of Coder 1.
- Table 3.11. Categorical analysis II of Coder 2.
- Table 3.12. Summary of the results of the acoustic and categorical analyses.
- Table 4.1. Forty rendaku “lovers”.
- Table 4.2. Results of rendaku.
- Table 5.1. The stem stimuli of the experiment.

- Table 5.2. Division of following sounds in the selected stimuli.
- Table 5.3. LOG lemma frequencies of the stimuli.
- Table 5.4. Age, gender, education level, and location of the subjects.
- Table 5.5. Mean height index of long vowel <ä> divided by the categorical variables.
- Table 5.6. Results of the pronunciation of long <ä> in the Alemannic variety of Standard German
- Table 6.1. Frequency counts of *koreksi*.
- Table 6.2. Results of mən- prefixation in Dutch loanwords in Indonesian.
- Table 7.1. Maps of the effect of geolinguistic factors on the pronunciation of Standard German long vowel <ä>.
- Table 7.2. The vowel height of long vowel <ä> in Standard German.
- Table 7.3. Results of the pronunciation of Standard German long vowel <ä> for East Middle German.
- Table 7.4. The effect of frequency on the pronunciation of long vowel <ä>.
- Table 7.5. The differences between long vowel <e> and <ä> in non pre-r context in Swabian-Alemannic in Switzerland, High Franconian and West Middle German.



---

# *Part I*

## *Introduction & Methodology*

---

*Tea is nought but this  
First you make the water boil  
Then infuse the tea  
Then you drink it properly  
That's all you need to know*

Sen no Rikyu 1522-1591

