The Identification of Bases in Morphological Paradigms

Adam C. Albright

University of California, Los Angeles

June, 2002

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Linguistics

UCLA Linguistics Department

Thesis committee:
Bruce Hayes, co-chair
Donca Steriade, co-chair
Carson Schütze
Brent Vine

Contents

List of Figures

List of Tables

ABSTRACT OF THE DISSERTATION

The Identification of Bases in Morphological Paradigms

by

Adam C. Albright
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 2002
Professor Bruce Hayes, Co-chair
Professor Donca Steriade, Co-chair

Many theories, in many domains of linguistics, assume that some members of morphological paradigms are more basic than others. Bases of paradigms are privileged in various ways: they may determine phonological properties of other forms, they may determine the direction of analogical changes, and so on. In this thesis, I propose that such effects are a result of the procedure by which learners seek to develop a grammar that allows them to project inflected forms as accurately and confidently as possible. I present a computationally implemented model of paradigm acquisition that attempts to use one form in the paradigm as the base to project the remaining forms, using stochastic morphological rules. I pursue two hypotheses about how this is done. The first is that learners are limited to selecting a single form as the base, and that the base must be a surface form from somewhere within the paradigm. Furthermore, the choice of base is global, meaning that the same slot must serve as the base for all lexical items. The second hypothesis is that learners select the base form that is maximally informative, in the sense that it preserves the most contrasts, and permits accurate productive generation of as many forms of as many words as possible.

As evidence for this approach, I analyze three cases in which an typologically marked form served as the base of a historical analogical change: Yiddish present tense paradigms (in which all forms were remodeled on the 1st sg), Latin noun paradigms (in which nominatives were remodeled on oblique forms), and Lakhota verbs (in which unsuffixed forms are being remodeled on suffixed forms). In each case, I show how the model correctly selects the base form, and also correctly predicts asymmetries in the direction of subsequent paradigmatic changes. I show that these asymmetries are not predicted by a more traditional model of underlying forms, in which learners compare all of the parts of the paradigm to construct abstract underlying representations that combine unpredictable information from multiple forms. Finally, I discuss possible extensions of this model to accommodate larger paradigms with multiple, local bases.

Acknowledgments

I would like to thank my committee members (Bruce Hayes, Donca Steriade, Carson Schütze and Brent Vine) for all of their patience, encouragement, and friendly criticism, and especially for agreeing to show up for a defense at 9am on a Sunday morning. First and foremost, I am thankful to Bruce Hayes, who has been a patient and wise teacher, an attentive advisor, and an enthusiastic collaborator during my time at UCLA. He has helped me more in my research and growth than many pages of acknowledgments could do justice to. I am also greatly indebted to Donca Steriade, for her wisdom and guidance, and for her knack for making hard problems seem so clear, at least until the time when I left her office. I am equally grateful to Carson Schütze for our many stimulating and challenging discussions in which he pushed me to take on hard issues, clarify my analysis, and see things from another angle. Finally, I am much obliged to Brent Vine for his thoughtful and insightful questions, for offering intriguing parallels and additional data, and for clearing up at least a few of my many misconceptions about Latin and Russian.

I would also like to thank all of the other members of the UCLA Linguistics Department who have made it such a vibrant and educational environment. I owe particular thanks to Colin Wilson for his discussions and advice, and to Pamela Munro for introducing me to Lakhota, and for generously providing me with her insights and the use of her Lakhota verb list. I am also especially grateful to Victoria Anderson, Marco Baroni, Rebecca Brown, Leston Buell, Ivano Caponigro, Taehong Cho, Christina Foreman, John Foreman, Matt Gordon, Amanda Jones, Sun-Ah Jun, Pat Keating, Sahyang Kim, Peter Ladefoged, Harold Torrence, Motoko Ueyama, Jie Zhang, Kie Zuraw, and all of the other members of the UCLA Phonetics Lab and Linguistics department (too numerous to name) who have made my years here stimulating and entertaining. In addition, I would like to thank my teachers during my undergraduate years at Cornell who instilled me with a love of phonology and historical linguistics—especially Abby Cohn, Wayne Harbert, and Jay Jasanoff.

Much of the data presented here would have been difficult or impossible for me to collect or make sense of on my own; I am particularly indebted to my Lakhota teacher Mary Rose Iron Teeth, for many hours patiently and cheerfully teaching me about her language. I am also extremely grateful to Angelo Mercado for his generous Latin help, and likewise to Jongho Jun, Taehong Cho, and Sahyang Kim for Korean, and to Katya Pertsova for Russian.

Various portions of this dissertation have benefited from questions and comments by audiences at UCLA, MIT, BU, the LSA (Chicago, January 2001), the ZAS Conference on Paradigm Uniformity (Berlin, March 2001), BLS 28 (Berkeley, February 2002), WCCFL 21 (Santa Cruz, April 2002), and the GLOW Phonological Acquisition Workshop (Utrecht, April 2002).

This dissertation work was supported in part by an NSF Graduate Fellowship, NSF grant

BCS-9910686, and a UCLA Dissertation Year Fellowship, all of which I gratefully acknowledge.

Writing a dissertation can be an ordeal, and it is made tolerable only by the unswerving support and encouragement of friends—even when you appear to have momentarily abandoned them. For this, I am eternally grateful to Moto and Marco, to Jie, Ivano, and Harold, to Scott, and most of all to Angelo. Last but not least, I am thankful to my parents, for their years of support and for giving me my first introduction to language.