

Linguistische Arbeiten

510

Herausgegeben von Klaus von Heusinger, Gereon Müller,
Ingo Plag, Beatrice Primus, Elisabeth Stark und Richard Wiese

Christian Uffmann

Vowel Epenthesis in Loanword Adaptation

Max Niemeyer Verlag
Tübingen 2007



Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN 978-3-484-30510-6 ISSN 0344-6727

© Max Niemeyer Verlag, Tübingen 2007

Ein Imprint der Walter de Gruyter GmbH & Co. KG

<http://www.niemeyer.de>

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Printed in Germany.

Gedruckt auf alterungsbeständigem Papier.

Druck und Einband: AZ Druck und Datentechnik GmbH, Kempten

Table of Contents

Acknowledgements	IX
1 Epenthetic Vowels in Loanwords.....	1
1.1 Selecting the Epenthetic Vowel	3
1.1.1 Default Segments	4
1.1.2 Vocalic Spreading	6
1.1.3 Consonantal Assimilation.....	7
1.2 Open Questions	9
1.2.1 The Interplay of Different Strategies.....	9
1.2.2 The Empirical Base	11
1.2.3 The Crosslinguistic Perspective	13
1.3 Conclusion.....	13
2 Theoretical Background	15
2.1 Optimality Theory.....	15
2.1.1 The Architecture of an Optimality-Theoretic Grammar.....	16
2.1.2 Classes of Constraints in Optimality Theory.....	18
2.1.3 Markedness, Faithfulness, and Epenthesis	19
2.1.4 Richness of the Base and Loanword Phonology	24
2.2 Segment Interaction in Optimality Theory.....	26
2.2.1 Strict Locality and Agreement	28
2.2.2 An Autosegmental Model of Optimality Theory	31
2.3 Strategies of Vowel Epenthesis	33
2.3.1 Default Epenthesis versus Feature Spreading	34
2.3.2 Do not LINK Dissimilar Segments.....	37
2.3.3 Do not SKIP Segments.....	40
2.3.4 A Typology of Epenthesis Strategies	41
2.4 Conclusion.....	44
3 An Empirical Analysis of Vowel Epenthesis in Shona	45
3.1 Resyllabifying Coda Consonants	49
3.1.1 Labials and Coronals.....	51
3.1.2 Dorsals	54
3.1.3 Liquids	56
3.1.4 Interim Summary and Generalization	58
3.2 Resyllabifying Onset Clusters.....	63
3.2.1 A First Statistical Analysis.....	64
3.2.2 Clusters of a Consonant and an Obstruent or Nasal	65
3.2.3 Clusters of a Consonant and a Glide	66
3.2.4 Clusters of a Consonant and a Liquid.....	69
3.2.5 Interim Summary and Generalization	73
3.3 Conclusion.....	78

4	An OT Account of Epenthesis in Shona.....	81
4.1	Syllable-Final Epenthesis: The Basic Patterns	81
4.1.1	Assimilation to the Consonant	81
4.1.2	Vowel Harmony after Dorsals.....	83
4.1.3	V-Place Spreading or Vocalic Spreading.....	86
4.1.4	Default Vowel Epenthesis.....	88
4.2	Harmony across Obstruents.....	89
4.2.1	Transparency and Sonority.....	90
4.2.2	Transparency and Vowel Quality.....	90
4.2.3	The Interaction of Vowel and Consonantal Quality.....	93
4.3	Epenthesis into Onsets	96
4.3.1	Directionality in Consonantal Spreading.....	97
4.3.2	Directionality in Vocalic Spreading	99
4.3.3	Epenthesis into Consonant-Glide Clusters	101
4.3.4	Interim Summary.....	103
4.4	Epenthesis in the Context of Liquids.....	103
4.4.1	Syllable-Final Epenthesis	104
4.4.2	Secondary Feature Specifications	106
4.4.3	Interim Summary.....	108
4.4.4	Consonant-Liquid Clusters.....	111
4.5	Conclusion.....	116
5	Shona Native and Loan Phonology	121
5.1	Transfer of Phonological Processes.....	121
5.1.1	Prosthetic /i-/	121
5.1.2	The Reversive Suffix.....	123
5.1.3	Height Harmony.....	124
5.1.4	Reduced Vowels.....	125
5.1.5	The Impact of Native Processes on Loanword Epenthesis.....	127
5.2	CV-Sequences and Stochastic Constraints	129
5.2.1	Analogical Modeling.....	129
5.2.2	The Database	131
5.2.3	The Analysis.....	132
5.3	Shona Vowel Specifications.....	135
5.3.1	Evidence from Loanwords Recapitulated	136
5.3.2	Phonetic Grounding.....	137
5.3.3	Underspecification in OT	139
5.3.4	Licensing and Native Shona Phonology.....	141
5.3.5	Summary	144
6	Vowel Paragoge in Sranan	145
6.1	Paragoge in Modern Sranan.....	146
6.1.1	The Statistical Analysis.....	147
6.1.1.1	Paragoge after /a/	149
6.1.1.2	Paragoge after /o/	151
6.1.2	The OT Analysis	153

6.1.2.1	Paragoge after /a/	155
6.1.2.2	Paragoge after /o/	156
6.1.3	Summary	158
6.2	The Diachronic Development of Paragoge in Sranan	159
6.2.1	Herlein (1718): Default Vowels	160
6.2.2	Van Dyk (c.1765): Emergent Vowel Harmony	161
6.2.2.1	Changing Default Vowels	164
6.2.2.2	Paragoge after Liquids	165
6.2.2.3	Paragoge after /a/	166
6.2.2.4	Interim Summary	167
6.2.3	Schumann (1783) and Nepveu (1770): Developing Vowel Harmony	167
6.2.3.1	Paragoge after /a/	170
6.2.3.2	Harmony and Sonority	173
6.2.3.3	Paragoge in Nepveu (1770)	174
6.3	Conclusion	175
6.3.1	The Development of Present-Day Patterns of Paragoge	176
6.3.2	A Comparison of Epenthesis in Shona and Sranan	179
7	The Crosslinguistic Perspective	183
7.1	Vowel Epenthesis in Kinyarwanda	184
7.1.1	Basic Epenthesis Patterns in Kinyarwanda	184
7.1.2	Epenthesis after Labials	186
7.1.3	Epenthesis after Coronals	188
7.1.4	Epenthesis after Dorsals	188
7.1.5	Epenthesis after Liquids	189
7.1.6	An Optimality-Theoretic Analysis	190
7.1.7	Concluding Remarks	194
7.2	Vowel Epenthesis in Samoan	195
7.2.1	Basic Epenthesis Patterns in Samoan	195
7.2.2	Epenthesis after Labials and Coronals	197
7.2.3	Epenthesis after Dorsals	198
7.2.4	Epenthesis after Liquids	199
7.2.5	Towards an Optimality-Theoretic Analysis	200
7.2.6	Summary	203
7.3	Epenthesis versus Deletion	204
7.3.1	The Optimality of Epenthesis	205
7.3.2	Word-Size Restrictors and the Optimality of Deletion	207
7.3.3	Epenthesis and Deletion in White Hmong	208
7.3.4	Triggers for Deletion	211
7.4	Default Epenthesis Revisited	213
7.4.1	Epenthesis of Placeless Vowels	213
7.4.2	Epenthesis in Developing Systems	216
7.5	Conclusion	217
8	Residual Issues	221
8.1	Phonology or Phonetics?	221

VIII

8.1.1	Phonetic Models of Adaptation.....	221
8.1.2	Against Phonetic Models of Adaptation.....	223
8.2	Alternative Models of Vowel Epenthesis.....	226
8.2.1	Perception-Based Models of Epenthesis	226
8.2.2	Strict Locality and Agreement	230
8.2.3	Rule-Based Models of Phonology.....	233
8.2.4	Summary	235
8.3	Conclusion and Outlook.....	236
References		239

Acknowledgements

This book is an abridged and revised version of my doctoral dissertation, which I submitted and defended at the Faculty of Foreign Languages of the University of Marburg. My biggest thanks thus go to my supervisors, most importantly the ever so patient Rüdiger Zimmermann, who helped in many respects, also outside phonological theory, by supporting my work, knowing how and where to tap resources, sending me abroad in the right moment, and giving me the freedom and backing to pursue my research. Great thanks also to my co-advisors, Ingo Plag and Richard Wiese. Ingo Plag introduced me to creole studies and to the world of academia by taking me along to my first conference. His constant criticism and encouragement during the writing process were also very helpful. Richard Wiese was always a good source of discussion and advice for all matters phonological. Thanks also for allowing me to participate in the German department phonology meetings, and thanks to all the people there, especially Wolfgang Kehrein and Birgit Alber; the discussions with you helped a lot. Special thanks to Birgit who not only stayed in touch and kept encouraging me after she had left for Trento but also for making it possible for me to stay and teach there for one semester.

In the English department, I have to thank my officemates with whom I could discuss many things linguistic (and less linguistic), first Sabine Lappe, then Marcus Callies, and to my colleagues and fellow graduate students. Big thanks to Wolfram Keller for making department life more fun and for widening my horizons in matters literary and Canadian, and for proof-reading. Particularly big thanks thus also to Jocelyn Keller who read through a bulk of the original draft. In addition, I want to thank my undergraduate teachers who raised my interest in linguistics, Jürgen Handke and particularly Michael Völtz, who spawned my interest in phonology and confirmed that loanwords were actually an interesting research topic.

I was fortunate to spend my first graduate year at Ohio State University, in the department of linguistics. I am greatly indebted to the DAAD for funding this year. At OSU, my biggest thanks go to David Odden for supporting my application and especially for being a great critical phonology teacher. Thanks also to Beth Hume for providing me with some crucial input that got the whole model I am presenting started, to Keith Johnson, who gave me a thorough and much-appreciated phonetic grounding, to Donald Winford, who opened up a new perspective on creole studies for me, to all the other teachers I interacted with and from whom I learned a thing or two, and to my fellow graduate students who were both scholarly inspiration and fun, especially all the co-phonologists and the 'Phonies' people.

I would also like to thank everybody at the University of Zimbabwe who made my stay there a very pleasant one, despite the political conditions and the odd armed robbery. Tatenda, Francis Matambirofa, who, as head of department, made things possible. A big thank-you also to Pedzisai Mashiri and Kennedy Chinyowa for being great Shona teachers. An especially big thank-you to Pedzisai for sharing his private collection of English loanwords in Shona with me. Thanks also to all the other linguists in the department who willingly discussed my ideas with me. Thanks to informant and friend Sipiwe Mutumwa and her family in Dangamvura township, Mutare.

I also want to thank people here at CASTL in Tromsø where I am based now. Special thanks to Curt Rice for discussing how to shorten the dissertation, but also to the other phonologists here who are a great source of inspiration and discussion, Sylvia Blaho, Patrik Bye, Martin Krämer, Ove Lorentz and Bruce Morén.

Thanks also to all the people I met at workshops and conferences who patiently listened to my ideas and provided me with valuable feedback. Thanks therefore to the audiences at the 1999 MOT workshop in phonology, the GGS 2000, HILP 5, NAPhC 2, OCP 1, MfM12 conferences, the regular and much-cherished DüköMarSie workshops, audiences at the University of Zimbabwe and the University of Southern California. Special thanks to Yvan Rose for allowing me to reanalyze his Kinyarwanda loanword corpus. Thanks also to all whose comments and remarks influenced me, especially Laura Downing, Emmanuel Dupoux, Janet Grijzenhout, Markus Hiller, Michael Kenstowicz, Bert Vaux, Rachel Walker and a few others which I might have forgotten at the moment (please feel included) for discussing my theories with me and providing me, perhaps unwittingly, with some important idea that had an impact on this book.

I also want to thank all my friends in Marburg who made the dissertation time worthwhile. Last but most importantly, my greatest thanks go to Anne Pieper for having been so supportive and patient with me even when I was on the verge of madness. Finally, I would like to thank my parents who made all this possible.

This work is dedicated to the ongoing struggle of the Zimbabwean people for freedom and independence.

1 Epenthetic Vowels in Loanwords

Vowel epenthesis is a frequently observed phenomenon in the adaptation of loanwords if the borrowing language has tighter phonotactic constraints than the donor language. Hence, languages with strict CV syllable structure epenthesize vowels in borrowed material if the donor language is more permissive with respect to allowed types of syllable structure, for example in permitting coda consonants or consonant clusters in onsets. Epenthesis thus occurs to avoid consonant clusters or consonants in coda position. Languages with less restrictive constraints on syllable structure, i.e. languages which allow syllables other than of the CV type alone, still epenthesize vowels as long as the structure is more restrictive than that of the donor language, for example, by allowing only certain consonants in coda position, disallowing clusters in codas or allowing only certain types of clusters (e.g. by allowing onset clusters consisting of an obstruent and a liquid but not clusters consisting of /s/ plus an obstruent). In (1), a few examples of English loans are provided from different languages to show that this is indeed a pervasive phenomenon (epenthetic vowels are in bold print).

(1) Epenthetic vowels in loanwords

Yoruba	k íláàsi	‘class’	(Akinlabi 1993)
Japanese	s utorai k u	‘strike’	(Park 1987)
Tswana	k eresemose	‘Christmas’	(Batibo 1995)
Shona	g irini	‘green’	(Uffmann 2001)
Samoan	s ikauti	‘scout’	(Cain 1986)
Fijian	s ipiini j i	‘spinach’	(Kenstowicz 2003)
Kikuyu	n girathi	‘glass’	(Mwihaki 2001)
Rennellese	k alapu	‘crab’	(Brasington 1978)
Haya	e sipurei	‘spray’	(Byarushengo 1976)
Luganda	s sukuru	‘school’	(Mosha 1971)

In all the above examples, we find epenthetic vowels in two different positions. First, to avoid onset clusters, vowels are inserted between the two consonants which occur in onset position in the etymon. Second, epenthetic vowels permit to syllabify consonants as onsets which are in coda position in the etymon. Frequently, these vowels occur word-finally as so-called paragodic vowels. Epenthesis thus avoids consonant clusters and coda consonants in languages where these are illicit.

Such epenthesis has been noted early in the literature, although comprehensive analyses of this phenomenon have remained rare to the present day. One early source of the creole language Sranan, the anonymous *Neger-Englische Grammatik* from 1863, already makes a surprisingly precise (although somewhat comical, from a present-day viewpoint) statement:

Die englischen Worte sind meist sehr verketzert, einige kaum wieder zu erkennen. [...] Gewöhnlich werden die aus fremden Sprachen übernommenen Worte neger-englisirt durch Hinzufügung eines Vokals am Ende des Wortes – fast immer desselben, der im Worte bereits

vorkommt. Auch in der Mitte eines Wortes schieben die Neger gern einen Vocal ein, wenn mehrere Consonanten auf einander folgen, welche sie schwer aussprechen können.

(Neger-Englische Grammatik 1863:1)¹

The anonymous author² thus confirms our opening claim, that vowel epenthesis occurs to avoid either word-final consonants or to break up consonant clusters. He also makes a statement about the quality of the epenthetic vowel, claiming that the epenthetic vowel is a copy of an underlying stem vowel.

Sadly, however, few analyses of vowel epenthesis in loanword adaptation have moved significantly beyond the findings of the anonymous author. Very often, a few blanket statements are made about the quality of the epenthetic vowel before the author proceeds to topics which they consider more worthwhile. The most important aim of this book thus is to fill this gap and to show that vowel epenthesis is a worthwhile research topic. We will show that patterns of vowel epenthesis are often quite complex, and in fact more complex than previous analyses would generally have it. The remainder of this chapter will therefore review the available literature and sketch an outline of the present state of research into vowel epenthesis.

Before moving to this point, another question needs to be answered first, however: Why do we find epenthesis at all? If only the well-formedness of syllables is at stake, this could be achieved via deletion as well. Across the literature, it has been noted time and again that vowel epenthesis is the most common strategy to satisfy syllable structure constraints. The other possible option, deletion, occurs much less frequently although it would also be a possible strategy to generate acceptable syllables in the borrowing language. Thus, it would be equally possible to borrow *class* as [ka], *strike* as [sai] or *green* as [gi], to name a few example words from (1) above. However, this is hardly ever the case. Across the board, epenthesis is found in loanword adaptation, as noted by many researchers on a wide range of languages (e.g. Wullschlägel 1856, Anonymous 1863, Smith 1977 on Sranan, Mosha 1971 on Luganda, Park 1987, Shinohara 1997, Katayama 1998 on Japanese, Pulleyblank 1988, Akinlabi 1993 on Yoruba, Kenstowicz 2003 on Fijian, among others). Batibo (1995), Rose (1995), Paradis (1996) and Plag & Uffmann (2000) also provide quantitative analyses for the languages they investigate. Batibo finds that deletion is found in only about 3 percent of cases in Tswana and Swahili. Rose (1995) finds only 1.4 percent deletions in his corpus of Kinyarwanda borrowings from French. Paradis (1996) gives similar numbers for her corpora of French loans in Moroccan Arabic, Kinyarwanda and Fula, and English loans in Quebec French (between 2.3 and 6.6 percent), and Plag & Uffmann (2000) report similarly low deletion rates from their investigation of diachronic Sranan corpora, ranging from zero to about 3 percent of types. Crosslinguistic evidence thus suggests that deletion is indeed a marginal phenomenon, compared to epenthesis.

¹ Translation: “The English words are very often mutilated, some hardly recognizable. [...] Usually the words taken from foreign languages are Negro-anglicized by adding a vowel at the end of the word – almost always the same occurring within the word already. In the middle of a word, the Negroes also like to insert a vowel, if several consonants follow one another, which they can hardly pronounce.”

² It is generally agreed today that the anonymous was H.R. Wullschlägel, also author of the *Neger-Englisches Wörterbuch* from 1856.

Paradis (1996) accounts for this phenomenon by invoking the Preservation Principle (see also Rose 1995, Paradis & Lacharité 1997), which holds that “segmental information is maximally preserved” (Paradis 1996: 511). The principle predicts that borrowers will try to preserve information found in the borrowed material as faithfully as possible, even at the expense of adding information (in the shape of epenthetic vowels), and indeed this is what seems to be found in loanword adaptation crosslinguistically.³ The Preservation Principle is posited largely on the basis of loanword data by Paradis and her colleagues; it is borne out by the data (as is repeatedly shown for a host of different languages), but seems like an ad-hoc stipulation; a principled explanation of why it exists is still a desideratum, especially if it is taken into account that epenthesis is generally considered universally marked compared to deletion (Vennemann 1988, Singh & Muysken 1995). This point is also taken up and discussed in Plag & Uffmann (2000), who show that epenthesis is a pervasive phenomenon not only in loanword adaptation but also in interlanguage phonology. Plag & Uffmann notice that the behavior of loanwords and second language phonology is puzzling if one subscribes to the view that deletion is a less marked process, as e.g. Singh and Muysken do, but they cannot provide a comprehensive explanation for the possibly unexpected behavior of loanwords. Is loanword adaptation thus a phenomenon *sui generis*?

For the moment, we will simply acknowledge that epenthesis is much more common than deletion but return to this point in §7.3 below when we will discuss some cases of languages where deletion actually is preferred over epenthesis, or where both strategies are attested, depending on context, and suggest a formal and functional explanation for the prevalence of epenthesis, arguing that deletion is only found when special high-ranked constraints render epenthesis suboptimal. The results will then be linked to research on epenthesis and deletion in language acquisition and language change. The question to be followed in the next section concerns the choice of the epenthetic vowel in those languages (the vast majority, as it were) which prefer epenthesis over deletion and to review the available literature on this topic.

1.1 Selecting the Epenthetic Vowel

If epenthesis is the general strategy in loanword adaptation in order to produce well-formed syllables in the borrowing language, the question then is which vowel precisely is inserted, in a given language as well as across languages. The literature is surprisingly silent about this issue; few analyses of loanword adaptation devote themselves to this question. In Shona, for example, which will be the main language of analysis in this book, there are quite a few works on loanword adaptation and interlanguage phonology (e.g. Baker 1947, Chimhundu 1979, 1983, Bernsten 1991, Pongweni 1991), but none of them addresses the problem of which vowel is inserted when syllable structure constraints trigger epenthesis, a

³ Paradis (1996) and Paradis & Lacharité (1997) do not exclude the possibility of deletions. However, they predict deletion to occur only if epenthesis is too costly, if it involves too many repairs (what they call the Threshold Principle).

problem which befalls other sketches and analyses of loanword phonologies as well, which prefer to discuss other phenomena like segment adaptation often only briefly touched upon epenthesis (e.g. Mosha 1971 on Luganda, Katayama 1998 on Japanese).

On the other hand, there is also a body of works which does discuss vowel epenthesis (in both loanword and native phonologies) and the problem of which vowel is epenthesized, although such analyses vary considerably in depth. By and large, three strategies which determine the quality of the epenthetic vowel have been identified in the literature. The epenthetic vowel can be a default vowel, one vowel which is invariably inserted across contexts. Alternatively, the vowel can be determined totally or in parts by the quality of an adjacent underlying vowel, as vowel copy or vowel harmony. Third, the quality of the preceding consonant can influence the choice of the epenthetic vowel. There are, however, very few investigations into this phenomenon that discuss the factors which favor one strategy over another, that try to delimit which strategies apply in different contexts and define where each strategy applies, attempting to give a comprehensive account of epenthesis in one language, supported by corpus analyses.

In outlining the three major strategies identified above, I will therefore introduce some of the more explicit approaches to the topic and their findings and then discuss several points of criticism which these approaches raise. It will be shown that there is a lamentable lack of comprehensive and thorough analyses of epenthesis patterns. Instead, many accounts rely on an impressionistic analysis of these patterns. In fact, the number of statistical analyses of epenthesis is close to negligible. Moreover, the vast majority of papers discuss only one selected language and do not compare the results obtained to findings in other languages, attempting a crosslinguistic typology of patterns. Hence, a principled analysis and discussion of vowel epenthesis in loanword phonology still is a desideratum in contemporary research, and this book will attempt to fill this gap by providing in-depth analyses of epenthetic vowels in several languages and by proposing a general typology of epenthesis patterns.

1.1.1 Default Segments

A number of papers and theses argue that default vowel insertion is a common strategy in loanword adaptation. A prime example of this is Japanese, for which a plethora of analyses demonstrate that /u/ is a default epenthetic vowel in loanword adaptation (e.g. Park 1986, Shinohara 1997, Katayama 1998); coda consonants and consonant clusters are resolved in adaptation by adding epenthetic /u/, unless the preceding consonant is /t,d/ in which case /o/ is epenthesized because /u/ would trigger affrication of /t,s/ to [ts,dz] (examples in (2); epenthetic segments are in bold print).

(2) Default vowel epenthesis in Japanese

fesutibaru	‘festival’	jiguzagu	‘zigzag’
disuku	‘disc’	furutaimu	‘full-time’
zippu koodo	‘zip code’	arubaito	‘job’ < German <i>Arbeit</i>

Default epenthesis is argued to occur across languages. Clements & Keyser (1983) show that in English, schwa is inserted into clusters which are prohibited in English, in words like *Gdansk*, *Phnom Penh* or *(Evel) Knievel* (the relatively permissive phonotactics of English render this a marginal phenomenon, though). Schwa is the default vowel in the native phonology of many other languages as well, e.g. German (Wiese 1998) and French (Noske 1982, Charette 1991). Pulleyblank (1988, 1998) argues that /i/ is the default epenthetic vowel in Yoruba, Byarushengo (1976) observes the same in Haya, and Kenstowicz (2003) makes a similar claim for Fijian loanword adaptation, while Abaglo & Archangeli (1989) suggest that /e/ is the default vowel in Gengbe, and Batibo (1995) argues for the status of /a/ as the default epenthetic segment in Tswana. According to Kitto & de Lacy (1999), default segmentalism is one of two possible epenthesis strategies (for the other strategy, see below). They cite a number of languages where /i/ is found as the default vowel in loanword epenthesis, e.g. Tongan and Cook Islands Maori, or in the native phonology, as in Lushootseed reduplication.

The question of what constitutes a default epenthetic vowel has received a number of different answers in phonological theory. One argument comes from phonetics: default vowels are shorter and thus less salient, an explanation which holds for Japanese where /u/ is frequently reduced, as shown by Beckman's (1982) instrumental analyses. In Japanese, /u/ is frequently shortened to almost zero duration. In addition, it can be devoiced, which further decreases its loudness and thus phonetic salience. In her P-map proposal, Steriade (2003) phonologizes such phonetic observations, proposing that epenthetic segments are those segments which are most confusable with zero because they are least perceptible in a given context. She shows that schwa is shorter in duration than other vowels, often unstressed and not stressable and of greater phonetic variability, which makes it less salient or perceptible and thus at the same time more confusable. Therefore, schwa is a common default epenthetic vowel, as in English, according to Steriade. Kenstowicz (2003) takes up Steriade's proposal and applies it to vowel epenthesis in Fijian where /i/ is the default epenthetic segment, according to Kenstowicz. He attributes the choice of /i/ to its relatively low perceptibility score. As /i/ is inherently shorter in duration than other vowels (it lacks the jaw movement necessary to produce low or mid vowels, Lehiste 1970), it is the vowel which is closest to zero and therefore selected for epenthesis. This (phonetic) argument is also mentioned in Byarushengo's (1976) analysis of Haya loanwords where the alleged default status of /i/ is also attributed to the relative shortness of the vowel (although Byarushengo admits that his hypothesis awaits experimental confirmation).

Other approaches are more phonological in nature. One explanation for the status of default vowels comes from theories of underspecification (Archangeli 1984, 1988, Pulleyblank 1988). In this theory, default epenthetic vowels are phonologically underspecified vowels. Pulleyblank (1988) therefore argues that /i/ is chosen as a default segment in Yoruba because phonologically, it is devoid of content – it is underspecified for all distinctive features. Abaglo & Archangeli (1989) make the same claim for Gengbe /e/. In theories of radical underspecification, features are filled in in the course of a phonological derivation via default rules. In Yoruba, /i/ does not bear any features underlyingly (the same argument holds mutatis mutandis for Gengbe /e/); features like [+high], [-back], [-round], [+ATR] are inserted later via rule application (on the notion of the phonetic grounding of such default rule applications, see Archangeli & Pulleyblank 1994).

The advent of Optimality Theory (Prince & Smolensky 1993, McCarthy & Prince 1993) has given rise to another possibility to explain the default status of some segments. At the heart of Optimality Theory (OT) there are markedness constraints, and default segments are simply the least marked segments, or in OT terms, segments which violate only the lowest-ranked markedness constraints. Epenthetic segments are then predictable by universal markedness relations which may interact with language-specific constraint rankings. Bermúdez-Otero & Börjars (2002) thus see vowel epenthesis as an emergence-of-the-unmarked effect since a default vowel is the least marked vowel; it does not violate any markedness constraints on vowels like *[+round] or *[+low]. Consequently, a typical default vowel, which satisfies markedness constraints on vowel realizations, is [i], [i] or [ə]. Lombardi (2003) builds upon similar observations, confirming that default vowels are the least marked vowels in the inventory. She assumes that central vowels, especially [ə] or [i], are universally least marked and thus chosen for epenthesis, followed by [i] which is an optimal epenthetic vowel in languages that do not have central vowels. Pulleyblank (1998) redefines the default status of Yoruba /i/ in terms of markedness relations which translate into faithfulness relations. A front high vowel then is the least marked vowel in a Yoruba-type 7-vowel system. High vowels are generally considered less marked than mid vowels (see also Beckman 1995, 1998), and front vowels are less marked than back vowels. In languages that have schwa, schwa will emerge as the default vowel because it is placeless and therefore does not violate any constraints on places of articulation.

Default vowel epenthesis thus is well-attested in the world's languages, in loanword adaptation as well as in the native phonology. Default vowels can be explained both phonetically or functionally (as the drive to insert a minimally salient, i.e. short and unobtrusive vowel) and phonologically, via underspecification theory or markedness relations. It has been argued, however, that the quality of the epenthetic vowel can also be derived through spreading, either from a vowel or from an adjacent consonant.

1.1.2 Vocalic Spreading

It has also been observed that the epenthetic vowel can be a full or partial copy of an adjacent underlying vowel (a preceding or following vowel). This observation is already found in the Sranan quote above which notes that the epenthesized vowel tends to be the same vowel as one already occurring in the word. This observation is confirmed by Smith (1977) who also argues that the epenthetic vowel in Sranan often is the same as the stem vowel, modulo some additional processes (which will be discussed below), a claim qualified by Plag & Uffmann (2000), which will also be put under scrutiny later (cf. §6).

Some researchers suggest that vowel copy or vowel harmony are the general epenthesis strategies in a number of languages. Paradis (1996) argues that the quality of the epenthetic vowel in Fula is determined by the nearest vowel or vocoid which is copied into the epenthetic slot. Kitto & de Lacy (1999) claim that copy is a second strategy in epenthesis, besides default segment insertion, and cite examples of languages where vowel copy is found, either as the only or the prevalent strategy, e.g. Selayarese, Awtuw and Winnebago, or as an additional strategy besides default segmentism, e.g. Cook Islands Maori. They also note intermediate processes in which part of the vowel is copied and the remainder is

inserted via default, i.e. vowel harmony, and cite Ponapean as an example where round harmony is found but the epenthetic segment is invariably [+high].

A number of works cite vowel copy or vowel harmony as one possible strategy among others. Rose (1995) finds that vowel harmony is one possible strategy in the adaptation of French loanwords in Kinyarwanda, but only across liquids (other epenthetic vowels will assimilate to the preceding consonant). Batibo (1995) argues that vowel harmony is the prevalent process found in Tswana epenthesis and a minor process in Swahili. Khumalo (1984) states that vowel copy is one minor epenthetic strategy in Zulu loanword adaptation (of English and Afrikaans words), alongside consonantal assimilation. Unlike Rose (1995), Batibo and Khumalo do not provide contexts, however, in which vowel copy or vowel harmony occur, that is contexts in which other strategies are blocked. Brasington (1978) also finds that vowel harmony or copy, which he somewhat confusingly dubs reduplication, is an important strategy in the adaptation of Rennellese loans.

Formally, such vowel harmony is commonly expressed as a spreading process. Rose (1995) uses autosegmental notation to express vowel harmony in Kinyarwanda borrowings, as do Batibo (1995) and Paradis (1996). Pre-autosegmental treatments (as in Smith 1977) may resort to SPE-type assimilation rules (Chomsky & Halle 1968), but the general theoretical implications are the same, in that the scope of one or several features is extended beyond the segment they are underlyingly associated with. Kitto & de Lacy (1999) follow a different route in their treatment of epenthetic vowels by proposing a theory within the framework of Optimality Theory. They claim that it is not the propagation of vocalic features which is responsible for vowel copy or harmony but correspondence between adjacent vocalic segments. Space does not permit giving a detailed explanation of their model here; the conception of spreading or assimilation as correspondence will not be followed in this book, however (for some arguments, see §2.2 and Uffmann 2005, to appear). There is no discussion of vowel harmony in epenthesis in Steriade's (2001) P-map model. It might be conceivable, though, to model it by assigning perceptibility scores to vowels in relation to other, neighboring vowels, such that identical adjacent vowels will have a greater confusability score and thus be rendered closer to zero than non-identical vowels. This issue will be taken up again in §8.2 when different formal approaches to epenthesis will be compared in the conclusion of this book. For the time being, Steriade's proposal will not feature prominently in the development of the argument but be evaluated briefly at the end.

1.1.3 Consonantal Assimilation

A third process is also mentioned quite frequently in the pertinent literature, assimilation of the epenthetic vowel to an adjacent (generally preceding) consonant. Here, the epenthetic vowel agrees with a neighboring consonant in its place of articulation. Probably the first to detect this process was Meinhof (1910) in his *Lautlehre der Bantusprachen*. He finds that

Umgekehrt entstehen Hülfsvokale bei Fremdworten oft zur Erleichterung der Aussprache. Sie richten sich in vielen Fällen nach den vorhergehenden Konsonanten und sind nach Labialen *u* bzw. *o*, nach *t, l, n* aber *i* bzw. *e*.

(Meinhof 1910: 14)⁴

Brasington (1978) asserts that such consonantal assimilation is an important strategy in epenthesis across languages, along with vowel harmony, that “the quality of epenthetic vowels [...] is regularly related, in absence of language specific pressures, either to that of preceding consonants [...] or alternatively, to that of neighbouring vowels” (Brasington 1978: 25). He finds that consonantal assimilation plays an important role in vowel epenthesis in Rennellese, alongside vowel copy. His observations are corroborated by a number of similar observations in other languages. The most widely cited example is labial attraction.

In labial attraction, a vowel assimilates in roundness to a preceding labial consonant. Hyman (1970) shows that /u/ is inserted in the context of labials in Nupe. Byarushengo (1976) notes that while /i/ is chosen as the default epenthetic vowel in Haya, /u/ is found following labials instead.⁵ Smith (1977) observes a similar behavior of epenthetic vowels in Sranan where vowel copy is generally found but /u/ is inserted after a labial consonant (but see §6 for a qualification of Smith’s statement). Akinlabi (1993) discusses loanwords in Yoruba and concludes that Pulleyblank’s (1988) default vowel analysis is too simplistic, since labial attraction (in his words, labial harmony) also plays an important role. Batibo (1995) shows that labial attraction figures prominently in Tswana and Swahili loanword adaptation, in a process he calls onset assimilation, whereby the epenthetic vowel assimilates in labiality (or non-labiality) to the preceding onset consonant: a round back vowel is inserted after a labial consonant, and a non-round front vowel is inserted after a non-labial consonant.⁶ Rose (1995) generalizes Batibo’s findings further in his discussion of French loanwords in Kinyarwanda. In his analysis, the epenthetic vowel generally assimilates in its place of articulation to the preceding consonant, such that a labial or round vowel will be inserted after a labial consonant and a coronal or front vowel will be inserted after a coronal consonant (dorsals are exempt and do not participate). Mwihiaki (2001) claims that in Kikuyu loans the preceding consonant is the only source that determines the quality of the inserted vowel: Labial /u/ is epenthesized following a labial consonant, and /i/ is epenthesized following coronals and dorsals, which Mwihiaki subsumes under the general

⁴ “On the other hand, supporting vowels frequently develop in foreign words to aid pronunciation. In many cases, they conform to the preceding consonant and are *u* or *o* after labials, but *i* or *e* after *t, l, n*” (my translation).

⁵ Byarushengo claims that /u/ is chosen after all non-coronals but only provides labials as evidence. In fact, the two examples involving a dorsal found in the paper are counterevidence to his claim since we find epenthetic /i/ there (*edesiki* from *desk* and *esaamusingi* from *something*).

⁶ Batibo uses a binary feature [labial], thus capturing not only cases of epenthesis of a round vowel after a labial consonant but also cases in which an unrounded, i.e. [-labial] vowel is inserted after a [-labial] consonant. His analysis therefore is more general than other analyses which only capture epenthesis of round vowels after labials but not the reverse. His analysis comes at the price of assuming a binary feature [labial], however, which is not uncontroversial since place features are commonly assumed to be unary or privative (see e.g. Clements & Hume 1995).

feature [+tongue] (cf. the Lingual node assumed in some models of feature organization). There is also epenthesis of /o/ following liquids; while she does not analyze how epenthesis of /o/ after a liquid could be understood as assimilation, it is still clear that the choice of the epenthetic vowel is determined by the preceding consonant (see also Byarushengo 1976 for a similar observation in Haya where liquids vocalize as /o/ which fits in with cross-linguistic observations about *l*-vocalization; see Johnson & Britain 2003 for an overview of phenomena involving the vocalization of /l/ as a back/round vowel or vocoid).

In sum, the effect of a (preceding) consonant on an epenthetic vowel is well-attested in many languages and a pertinent feature of loanword adaptation. This effect is commonly understood as assimilation and consequently expressed as such, either in the shape of rules in the generative tradition (e.g. Hyman 1970, Muihaki 2001) or as spreading rules in autosegmental frameworks (e.g. Batibo 1995, Rose 1995). Again, the exceptional analysis comes from Kitto & de Lacy (1999) who treat assimilation as correspondence in their OT account of vowel epenthesis. While they do not discuss consonantal assimilation in detail (but note its existence and quote Southeastern Pomo), they also seem to view it as correspondence, not spreading, a view which we reject in this book.

1.2 Open Questions

The question of which vowel can be chosen in epenthesis thus already received attention, and there is sufficient evidence for three different strategies, default epenthesis, vowel copy or harmony and consonantal assimilation. However, there are also still significant gaps in our understanding of loanword epenthesis which relate to three distinct areas in which research has so far been less conclusive. First, it is still unclear how the different strategies can interact. Do we find different strategies within one language? If so, can we predict when which strategy is chosen or do these strategies occur randomly side by side, across contexts? Second, the empirical base over which such generalizations are made is often unclear. In-depth investigations are rare compared to relatively impressionistic statements. Third, many investigations deal with single languages only. A crosslinguistic perspective, which can yield a typology of epenthesis strategies, is still lacking. This section will discuss the three problematic points in greater detail.

1.2.1 The Interplay of Different Strategies

Of those authors who suggest that several strategies may be active simultaneously in one language (many name only one process, as we have seen above), few are explicit about the contexts in which a given process may apply. This is particularly the case if vowel harmony and consonantal assimilation are both mentioned as productive strategies in that language. For example, Khumalo (1984) mentions labial attraction as a general strategy in the adaptation of English and Afrikaans loans in Zulu, such that /u/ is epenthesized following a labial and /i/ is epenthesized elsewhere. He also notes occurrences of vowel copy but remains silent about whether vowel copy and consonantal assimilation occur in the same

contexts or whether it is possible to make a more precise statement about the applicability of each process. The same problem befalls Batibo's (1995) analysis of loans in Tswana and Swahili. He mentions three general strategies that can apply when a vowel is epenthesized, consonantal assimilation, vowel harmony and default insertion of a low vowel, and shows that they apply at different frequencies in Swahili and Tswana. However, he does not say anything about the contexts in which each strategy applies (except for one statement that vowel harmony is probably more likely to occur word-finally in Tswana). Instead, he seems to suggest that each strategy can apply in any context, if only at different rates. The same holds for Brasington's (1978) analysis of Rennellese loans. Brasington identifies consonantal assimilation and vowel harmony as the two most important strategies in the selection of the epenthetic vowel and provides a statistical analysis of epenthesis patterns to show that both are significant. However, he cannot pin down where each process applies. He finds that vowel harmony is more likely to occur in onset clusters than in codas (the reverse of what Batibo finds) but generally, both processes seem to coexist side by side, across contexts.

One of the few analyses that is more explicit is Rose's (1995) discussion of Kinyarwanda borrowings from French. Rose also uses statistics and a large Kinyarwanda loanword corpus. He looks at epenthetic vowels in different contexts and finds that vowel harmony, consonantal spreading and default insertion are found in different contexts, consonantal assimilation being the general case and vowel harmony and default insertion occurring in clearly defined domains: Vowel harmony is found after liquids, which do not spread, and default vowel insertion is found after dorsals, which also do not spread but which, in addition, are also not transparent for vowels and hence block harmony under his analysis. Generally, however, such precise statements about the contexts in which different processes apply are rare in the available literature.

There is more precision in those analyses which claim that a default vowel is generally inserted in the language under investigation, but that one or two additional processes may additionally apply (the above examples were only concerned with vowel harmony and consonantal assimilation as rival processes, under the exclusion of default insertion). The context of those additional processes then is generally quite well described, probably because default insertion of one invariant vowel can be seen as the general or elsewhere case, and deviations from this invariant process can then be pinned down more exactly than in the interaction of two processes which both yield variable epenthetic quality. Thus, Byarushengo describes Haya epenthesis as a case of default epenthesis of /i/, unless labial attraction forces consonantal assimilation. A similar analysis is suggested for Yoruba by Akinlabi (1993). Kitto & de Lacy (1999) describe Cook Islands Maori as a case where /i/ is generally epenthesized but vowel copy is found across /r/. Smith (1977) analyzes epenthesis of /i/ in Sranan as the elsewhere condition, while vowel harmony and consonantal assimilation as labial attraction occur in more specific contexts.

What all the above examples show, however, is that there is a lamentable lack of in-depth analyses that discuss the domains in which rival epenthetic processes can occur, especially if they concern both vocalic and consonantal spreading and not just a small set of clearly defined exceptions from default segment insertion. The exception is Rose (1995), who discusses French loanwords, though (all other works discussing the borrowing of English words into a target language). It is unclear whether these processes occur in identical phonological contexts or whether they can be allocated to different contexts, so

that the choice of epenthetic strategy is predictable. It is therefore one of the major aims of this book to have a closer look at such rival processes and to establish whether they apply randomly across contexts or whether their occurrence is predictable. It will be argued in the course of this work that the latter is the case. Many languages display alternations between local, consonantal assimilation and vocalic assimilation or spreading which can be described unambiguously and with high predictive power, if a thorough empirical exploration of loanword data is done first. This point, thorough empirical exploration, leads me to the second point of criticism, viz. the lack of such analyses in the literature on vowel epenthesis in loanword phonology.

1.2.2 The Empirical Base

It is unclear in many analyses of vowel epenthesis in loanwords what exactly the empirical base is over which the generalizations are made which have been the subject of discussion so far. Many accounts remain conspicuously silent about the size of the loanword corpus used, which may lead one to the suspicion that the generalizations are based on introspection and on impressionistic interpolations rather than on a solid empirical basis. If this is the case, one should also view the generalizations themselves with suspicion. Is it really the case that language X invariably epenthesizes one default vowel, or has the researcher perhaps overlooked data which show that in certain, maybe less frequent contexts different strategies are found? The same holds for analyses which claim that there is default epenthesis but that in only one specific context another process may apply. Here, too, it may be found that the pattern really is more complex and that what has been regarded as default epenthesis might in fact not be default epenthesis but simply the frequent occurrence of one pattern that should be explained differently.

Let us consider one example to clarify the point. Byarushengo finds that in Haya /i/ is frequently found as the epenthetic vowel. He therefore assigns default status to /i/. He also notes, however, that there are cases in which /u/ is inserted, but after a labial consonant. He thus invokes a second process, labial attraction, to account for the deviant pattern.⁷ What, then, is the problem with this analysis? Note that the pattern that Mwihaki describes for Kikuyu looks conspicuously like the Haya pattern. She comes to a different conclusion, though. To her, all epenthesis can be understood as consonantal assimilation, in which a round or labial vowel is inserted after all labial consonants and a non-round vowel is inserted after all non-labial or lingual consonants, an analysis also found in Batibo's (1995) discussion of Tswana and Swahili loans and Khumalo's (1984) account of Zulu borrowings.⁸ Two points can be brought forward in favor of Mwihaki's (and Batibo's and

⁷ Byarushengo does not make labial attraction responsible for this process but a more general [grave] attraction, claiming that non-coronals in general trigger insertion of /u/. The lack of examples provided (and the existence of counterexamples) leads me to recast Byarushengo's analysis as a (more frequently attested) case of labial attraction.

⁸ The pattern found in Kinyarwanda is also very similar. Rose (1995), however, comes to yet a different conclusion, claiming that the insertion of /i/ after dorsals is not a case of consonantal

Khumalo's) analysis and against Byarushengo's analysis. First, it simplifies the analysis by assuming just one process, assimilation to the consonant with respect to labiality, instead of positing two processes, default epenthesis and assimilation. Second, thorough investigation of the data may help to find additional processes which aid in the decision. Mwihiaki finds that consonantal assimilation also produces a third pattern, insertion of /o/ after liquids, which supports her analysis of Kikuyu epenthesis as a case of assimilation to the preceding consonant. Then, the high frequency of /i/ as an epenthetic vowel does not follow from its default status but only from the high frequency in which spreading of [-labial] applies.

Despite the importance of a good empirical base, very few works thus seem to rely on one.⁹ Some give at least raw counts of epenthetic vowels to indicate frequencies, as Kenstowicz (2003) does to back up claims about default segmentism or deletion vs. epenthesis in specific contexts. Batibo (1995) provides percentages of the different adaptation processes he finds in Tswana and Swahili. Both thus show that at least they use a sizeable loanword corpus for some basic calculations. More elaborate statistical explorations of loanword corpora are rare, though. Brasington (1978) uses some statistics in his discussion of Rennellese loans, finding that both vowel copy and consonantal assimilation are significantly frequently used strategies in crosstabulations and subsequent chi-square significance tests. While he can show that the quality of an adjacent vowel or consonant has a significant impact on the selection of the epenthetic vowel, he fails to provide rules or generalizations, however, when each process occurs, implicitly suggesting instead that both may occur randomly in a given context. Again, a notable exception is Rose's (1995) discussion of French loans in Kinyarwanda. Rose uses a large corpus of 756 borrowings which he divides up into smaller subsets, according to different contexts (preceding obstruent vs. preceding liquid, cluster types, etc.) and then does frequency counts on the different subcorpora. Again, he does not use sophisticated statistics but at least he draws upon a solid empirical base and systematic heuristics to arrive at the complex pattern of vowel epenthesis that he finds, a method also found in the analyses of Paradis (1996) and Paradis & Lacharité (1997), although they do not discuss the choice of the epenthetic vowel.

In sum, the empirical base of discussions of loanword adaptation seems to be fairly weak in many cases. In addition, there is a general lack in the use of sophisticated statistical methods which can aid the researcher in the exploration of a large-size loanword corpus. This book will thus try to remedy this situation by providing in-depth statistical explorations of large-size loanword corpora. For example, the Shona loanword corpus, which will be at the heart of this thesis (§§3-5), contains 1709 types, some of which also have variants, such that the total number of types used well exceeds 2000. Somewhat

assimilation but of default epenthesis instead. This issue will be taken up in chapters 3 and 4 again, where a similar pattern found in Shona is discussed and Rose's analysis is largely confirmed.

⁹ Again, it should be noted that this claim is only based on an informed guess. The absence of discussions about corpus size and exploration techniques does not mean that they are not used. I therefore apologize for what some might feel are unwarranted accusations. The apparent superficiality of some analyses, however, especially in the light of the patterns that emerge in the in-depth analyses provided in this work, suggest that many analyses are in fact based on impressionistic observations and introspection rather than hard data.

smaller-size corpora of Sranan (several diachronic corpora), Samoan and Kinyarwanda, each still consisting of several hundred items, will also be explored using the same methods, before the findings will enter a formal, theoretical modeling of vowel epenthesis. It will be shown that in each of the corpora the observed pattern is more complex than most of the available literature would have it. In all four languages investigated, default insertion, vowel harmony and consonantal assimilation play an important role, in an intricate interaction of processes. In addition, it will also be shown that the manner of articulation of a preceding consonant can play a role in the choice of the epenthetic vowel, a hitherto unobserved behavior. The use of large corpora, in combination with statistical tools, can thus be more than gizmoism or analysis for analysis's sake. It can help us to detect patterns that mere introspection would overlook.

1.2.3 The Crosslinguistic Perspective

A third point of criticism concerns the lack of cross-linguistic comparisons found in the loanword literature. Shinohara (1997) and Katayama (1998) discuss Japanese, Byarushengo (1976) discusses Haya, Mwihi (2001) discusses Kikuyu, Kenstowicz (2003) analyzes Fijian, and so forth. Batibo (1995) is already an exception because he compares two languages, Tswana and Swahili, finding that both languages use the same strategies but at different frequencies. While vowel harmony is predominantly found in Tswana (followed by consonantal assimilation and default insertion, in this order), Swahili uses consonantal assimilation as the major strategy to adapt English loans to the native phonology; vowel harmony and default epenthesis are only marginally found in Swahili. Still, Batibo only states his findings but does not develop a broader typological perspective.

Consequently, this typological perspective still is very much of a desideratum in loanword phonology. There have been no attempts to synthesize the different findings into a unified theory of vowel epenthesis in order to show which processes are most frequently found and to try to find explanations for why certain processes are found in certain languages in certain environments. To my knowledge, there is only one work which deals with this issue and proposes an optimality-theoretic analysis of different epenthesis strategies, Kitto & de Lacy's (1999) paper on vowel epenthesis. They try to provide a unified account of assimilatory epenthesis (most notably vowel copy or vowel harmony) and default segmentism in a correspondence-based framework. However, their discussion of different epenthesis patterns remains somewhat superficial, especially in the light of the empirical findings presented in this thesis. A more general crosslinguistic typology and optimality-theoretic analysis will be proposed in the following chapters.

1.3 Conclusion

Despite our knowledge about why epenthesis occurs in loanwords (tighter phonotactic constraints in the borrowing languages which may disallow consonant clusters or coda consonants), we are still far from a comprehensive theory of epenthesis, especially with

respect to the question of which vowel is chosen in epenthesis. Starting out with the observation that epenthesis, rather than deletion, is a pervasive phenomenon in loanword adaptation, the available literature on epenthetic vowels was then reviewed, showing that three epenthetic processes are attested, default segment insertion, vowel harmony and consonantal assimilation. However, we still know fairly little about how these processes are distributed across languages as well as within a language, across contexts.

This work will try to fill this gap by providing detailed analyses of loanword corpora. Both empirical-statistical and formal analyses will be conducted in order to answer the following questions: What exactly are the observable patterns of epenthesis in a given language and how can this feed into a general theory of epenthetic vowels? The above review of recent literature on this topic demonstrates that these questions are still largely unanswered. Too many analyses have relied on small corpora which are explored in a somewhat superficial manner, while those – relatively few – analyses which rely on a broader empirical basis show time and again that the actual patterns of epenthesis are much more complex than many assume. It is therefore necessary that a thorough investigation of epenthesis patterns is carried out to establish what the patterns really are. The remainder of this book will be devoted this issue. We will introduce the theoretical framework assumed in chapter 2, a combination of Optimality Theory (McCarthy & Prince 1993, Prince & Smolensky 1993) and Feature Geometry (Clements 1991, Hume 1990, 1992, Clements & Hume 1995). The largest part of the book will deal with an in-depth analysis of vowel epenthesis in English loanwords in Shona. Chapter 3 will provide a detailed statistical analysis of epenthesis patterns and their feature-geometric representations, followed by a comprehensive optimality-theoretic analysis of these patterns in chapter 4 and a discussion of how these patterns interact with and can be motivated from Shona native phonology in chapter 5. The subsequent chapters will broaden the empirical base. Chapter 6 is a discussion of epenthesis in Sranan, both from a synchronic and a diachronic perspective. Additional analyses of Kinyarwanda and Samoan will be provided in chapter 7, along with a discussion of the crosslinguistic implications of the findings. Chapter 8 concludes and relates this book to current debates in the loanword adaptation literature.

2 Theoretical Background

This chapter introduces the theoretical frameworks that will be used, Optimality Theory and Feature Geometry. The model proposed in this work will combine both theories in a novel approach in which constraints are sensitive to feature-geometric representations. It thus re-emphasizes the importance of representations in phonological theory, rejecting the claim that output constraints alone can exhaustively account for all phonological patterns.

We will begin with a brief introduction to Optimality Theory in §2.1 and show how epenthesis in loanword adaptation is modeled in this framework before outlining the specific model of segment interaction we will assume in this book, a theory of constraints on autosegmental representations, and briefly defending this model against other approaches within the OT framework (§2.2). For a more detailed critique of alternative models and a defense of representationally strong OT, see Uffmann (2004, 2005, in press). The model we suggest will then be detailed in §2.3, presenting a first preliminary typology of epenthesis strategies in loanword adaptation, building upon the observations collected in chapter 1.

2.1 Optimality Theory

The advent of Optimality Theory (OT; Prince & Smolensky 1993, McCarthy & Prince 1993) meant a major paradigm shift for phonological theory. In traditional, SPE-type theories of phonology (Chomsky & Halle 1968), an underlying representation is transformed into a surface form via the stepwise application of rules in a serial derivation. This view of phonology changes radically in OT. Instead of assuming serialism, there is only one level of derivation on which constraints on outputs (not rules) are evaluated in parallel. Two things stand out in particular: first, the focus on constraints and second, the output-orientedness of OT – it is no longer necessary to establish the exact shape of an underlying representation first (more on this below in §2.1.4 when Richness of the Base and its implications are discussed). The paradigm shift is thus a double shift – not only from rules to constraints, but also from an input-based to an output-based view of phonology. The notion of constraints is of course not new in phonological theory; from the beginnings of Generative Phonology, they were present as additional conditions on the well-formedness of phonological representations, e.g. as morpheme structure conditions (Chomsky & Halle 1968) which hold on the level of underlying forms, or in order to explain the apparent ‘conspiracy’ of rules (Kisseberth 1970) which converge on similar output shapes. OT, however, removes the rule component from the theory altogether, keeping only constraints, and additionally limits these to constraints on outputs alone.

These constraints are characterized by three important properties: they are universal, they are violable, and they are ranked. The universality of constraints means that all languages share the same set of constraints CON, that the differences between languages

are not due to the existence of different constraints.¹ The locus of variation is somewhere else in OT - it lies in the stipulation that constraints are ranked with respect to each other. Higher-ranked constraints take preference over lower-ranked constraints (the principle of strict domination). The world's languages therefore differ in how they rank their constraints, which degrees of importance they assign to the individual constraints. This enables the researcher to devise factorial typologies which should match typological observations about actual grammars: Different rankings yield different grammars, and the sum of potential rerankings makes a prediction about which grammars should be possible (i.e. attestable) or impossible. The issue of factorial typologies will be taken up again throughout the book, and towards the end of this chapter a first rough typology will be proposed for types of epenthetic vowels. In addition, constraints are violable, a hypothesis which departs from traditional phonological theories where constraints, usually in the shape of well-formedness conditions, are considered inviolable. Violation is minimal, though: The output form (the optimal form) violates as few high-ranked constraints as possible.

2.1.1 The Architecture of an Optimality-Theoretic Grammar

Apart from CON, the universal constraint set, an OT grammar consists of two further important elements, a generator function GEN and an evaluator function EVAL. From an input (the underlying form), GEN generates a set of candidates (possible outputs) which is then passed to the evaluator function EVAL. The candidate set is generated by freely performing operations on the phonological string provided by the input ("freedom of analysis", McCarthy & Prince 1993: 21); the set is thus potentially non-finite, since every operation or type of operation may in principle be performed an infinite number of times. This set of candidates is then evaluated by EVAL, which contains CON, and the most harmonic candidate, that is the candidate incurring the fewest violations of high-ranked constraints, is selected as the output candidate. The diagram in (3), adapted from Archangeli (1997), shows graphically how the components of an OT grammar interact.

¹ It is generally assumed that the set of constraints is provided by Universal Grammar, that constraints are innate. More recently, however, the idea was raised that constraints may be learnable (Hayes 1996, Boersma 1998, Ellison 2000, Curtin 2001, Bermúdez-Otero & Börjars 2002), yet universal because they are functionally grounded. This book does not take sides in the argument, as the generalization that the same set of constraints is employed in all languages is not jeopardized by any of the theories.