English stress preservation and Stratal Optimality Theory

Sarah Collie

Thesis submitted for the degree of PhD by Research

The University of Edinburgh

2007

Abstract

Since Chomsky & Halle (1968), English stress preservation – *or**í**ginal → or**ì**ginálity*, ***ób**vious → **ób**viousness* – has been important in generative discussions of morphophonological interaction. This thesis carries out empirical investigations into English stress preservation, and uses their results to argue for a particular version of Optimality Theory: Stratal Optimality Theory ('Stratal OT') (Kiparsky, 1998a, 2000, 2003a; Bermúdez-Otero, 1999, 2003, in preparation). In particular, the version of Stratal OT proposed in Bermúdez-Otero (in preparation) and Bermúdez-Otero & McMahon (2006) is supported.

The empirical investigations focus upon the type of preservation where preserved stress is subordinated in the preserving word ('weak preservation'): e.g. *or**í**ginal → or**ì**ginálity*; ***àn**t**í**cipate → **ăn**t**ì**cipátion*. Evidence for the existence of weak preservation is presented. However, it is also shown that weak preservation is not consistently successful, but that it is, rather, probabilistically dependent upon word frequency. This result is expected in light of work like Hay (2003), where it is proposed that word frequency affects the strength of relationships between words: stress preservation is an indicator of such a relationship.

Stratal OT can handle the existence of English stress preservation: by incorporating the cyclic interaction between morphological and phonological modules proposed in Lexical Phonology and Morphology ('LPM'), Stratal OT has the intrinsic serialism which is necessary to predict a phenomenon like English stress preservation. It is shown that the same cannot be said for those of models of OT which attempt to handle preservation while avoiding such serialism, notably, Benua (1997).

Bermúdez-Otero's (in preparation) proposal of 'fake cyclicity' for the first stratum in Stratal OT can capture weak preservation's probabilistic dependence upon word frequency. Fake cyclicity rejects the cycle which has previously been proposed to handle weak stress preservation, in LPM and elsewhere; instead, fake cyclicity proposes that weak preservation is a result of blocking among stored lexical entries. Blocking is independently established as a psycholinguistic phenomenon that is probabilistically dependent upon word frequency; in contrast, the cycle is not a

probabilistic mechanism, and so can only handle instances of stress preservation failure by stipulation.

Acknowledgements

This thesis could not have been completed without the assistance of my supervisors, Heinz Giegerich and Patrick Honeybone. Their formidable phonological and professional expertise has been invaluable, as has their enthusiastic encouragement, tolerance, patience, sound advice and good humour.

Ricardo Bermúdez-Otero has also played a key role in the development of the following work by providing me with useful references, copies of manuscripts, and prompt, extremely thorough replies to late-night bursts of emails. I develop many of Ricardo's ideas in the latter chapters of this thesis; needless to say, responsibility for any errors or weaknesses in the analyses and argumentation is mine alone.

Many other people also deserve thanks. My work has benefited greatly from feedback at several conferences, and I am very grateful to my audiences at the Third Cambridge Postgraduate Conference in Language Research (Cambridge, 2005), Directions in English Language Studies (Manchester, 2006), On Linguistic Interfaces (Belfast, 2007), and the Second International Conference on the Linguistics of Contemporary English (Toulouse, 2007). With respect to events here at Edinburgh, I am appreciative of the feedback I have received from my colleagues at the Linguistics and English Language Postgraduate Conference and at the English Language Research Group. Additionally, both Nicolas Ballier and Michael Hammond have been helpful email correspondents.

I have consulted several members of the Department of Linguistics and English Language at Edinburgh at one time or another: Michael Bennett, Rob Clark, Bob Ladd, April McMahon, Satsuki Nakai, Mits Ota and Alice Turk have all helped this thesis along in some way. My fellow postgrads in 8 Buccleuch Place have been good company and always ready for a chat.

Iain and my family and friends must be thanked for their support and for providing me with distractions.

I would like to acknowledge the financial support of the Carnegie Trust for the Universities of Scotland, without which I could not have undertaken this work.

I confirm that this is my own work, written by me, and that it has not been submitted for any other degree except as specified.

Contents

Index of constraints

Due to variation in the practices of different researchers, constraints with different names may have the same definition; where this is the case, it is indicated here. Constraints which are innovations by the present work are indicated here by the use of boldface type.

Faithfulness constraints

Prosodic markedness constraints

Featural markedness constraints

Alignment constraints

Other phonological constraints

List of abbreviations and symbols

Acute accent (´)	Main stress
Grave accent (`)	Stress with secondary prominence
˘	Stress with tertiary prominence
ˈ	Stress, level of which is unimportant
< >	Delimit an extrametrical prosodic constituent
( )	Delimit a foot
*T**ì**c**ò**nderóga*	**In bold**, two stresses of non-primary prominence where the distinction between secondary and tertiary prominence is irrelevant
σ	Syllable
H	Heavy syllable
L	Light syllable
#	Word boundary
[ ]	**In chapters 4 and 5,** the dictionary pronunciation
/ /	**In chapters 4 and 5,** an underlying representation deduced from a dictionary pronunciation

Introduction

0.1 Introduction to the thesis

This thesis is about the nature of English stress preservation, and how this phenomenon should be handled theoretically.

The term 'stress preservation' denotes a situation whereby a complex word mirrors the stress pattern of a word embedded within it: for example, in English, *or**ì**ginálity* (*or**í**ginal*), ***ób**viousness* (***ób**vious*); in German (following Alber, 1998), *ùnivers**à**litä́t* 'universality' (*ùnivers**á**l* 'universal'), *ìrration**à**litä́t* 'irrationality' (*ìrration**á**l* 'irrational'). Stress preservation is, therefore, a phonological indicator of a relationship between words.

English stress preservation has been central to generative accounts of the phonology of English since the publication of Chomsky & Halle's *Sound Pattern of English* some forty years ago, where preservation was a showpiece phenomenon for one of this work's central theoretical innovations – the cycle. Continued interest in stress preservation can be attributed to its relevance to two key and interrelated areas of discussion in Generative Phonology: the nature of the interface between morphology and phonology, and the issue of phonological opacity. This thesis continues to explore these areas of theoretical debate in its advocation of Bermúdez-Otero's (in preparation) model of Stratal OT.

The weak stress preservation that is the focus of this thesis is a sub-type of English stress preservation.[1] With weak stress preservation, the stress pattern of the embedded word is subordinated in the preserving, embedding word, as is shown by the examples in (1):

[1] The terms 'weak preservation' and 'strong preservation' that are used here were first proposed by Burzio (1994). The following symbols are used to denote different levels of prominence among stressed syllables: ' ´ ' for primary prominence; ' ` ' for secondary prominence; ' ˘ ' for tertiary prominence. The argument for the recognition of tertiary prominence is given in §1.5.1.

(1) English weak preservation

Embedding word	Embedded word
(a) or**ì**ginálity	or**í**ginal
(b) s**ĕn**s**à**tionálity	s**èn**s**á**tional
(c) cònd**ĕn**sátion	cond**én**se

Weak stress preservation causes the stress patterns of morphologically complex words to differ from those of phonotactically equivalent morphologically simple words. This can be seen in (2):

(2) Contrast between complex and simple words

Complex word	Simple word
(a) or**ì**ginálity	**à**bracadábra
(b) s**ĕn**s**à**tionálity	**Tì**c**òn**deróga~**Tì**c**ŏn**deróga
(c) cònd**ĕn**sátion	Àrg**en**tína

Contrasts like those in (2) provide evidence for the existence of stress preservation.

Weak stress preservation will be further sub-categorised in the following chapters. Type (1a), exemplified by *orìginálity* (*oríginal*), and (1b), exemplified by *sĕnsàtionálity* (*sènsátional*), will both be classified as types of left-edge stress preservation, as the preservation occurs towards the left edge of the embedding word, two or more syllables before the main stress. Type (1c), where preservation is argued to occur on the pre-tonic syllable of the embedding word, as in *còndĕnsátion* (*condénse*), will be treated separately, and, indeed, will be argued not to be a type of stress preservation at all in the following chapters. The type of stress subordination seen under compounding, e.g. ***làw** degree lánguage requirements* (***láw** degree*), resembles the within-word preservation seen in (1), but Liberman & Prince (1977) propose that compound and phrasal stress should be handled quite differently from within-word preservation, hence compound stress is not examined in this thesis.

In English, there is also stress preservation where the prominence of the embedded word's stress is not subordinated in the embedding word – strong preservation. Strong preservation occurs with 'stress neutral' suffixation, and can be seen in examples like those in (3):

(3) English strong preservation

Embedding word	Embedded word
h**ópe**lessness	h**ópe**less
n**á**tionalising	n**á**tionalise
óbviousness	**ób**vious

Strong preservation resembles the free stress that is seen in languages such as Russian, where, as described by Hyman (1977: 39), the position of the main stress of a word is fixed not in relation to the word edge, but, rather, with respect to a particular syllable in the base morpheme. As with weak stress preservation, strong stress preservation creates a situation whereby a word's morphological structure affects its stress pattern. In English, primary stress is usually assigned to one of the final three syllables of a word, the precise syllable being determined by syllable weight and lexical category distinctions: e.g., *intervéne*$_V$, *húrry*$_V$, *pérvert*$_N$, *América*$_N$. Strong stress preservation can cause this three-syllable generalisation to be violated, as in *nátionalising* and *óbviousness* in (3). Strong preservation is not the subject of empirical investigation in this thesis: it will be seen from the following chapters that the phenomenon of English weak stress preservation is controversial enough to consume a thesis with its discussion. However, this is not to say that strong preservation is a completely cut-and-dry phenomenon, and a particular area for future exploration is noted in the thesis' conclusion.

Up until now, stress preservation has simply been assumed to occur in English, with no support from any serious and extensive empirical investigation. In light of the theoretical debate arising from Optimality Theory (OT) (Prince & Smolensky, 2004 [1993]) over the last decade or so, it is important that this situation is resolved. This is because, although OT is generally recognised as being successful in its handling of prosody, a phenomenon like stress preservation is problematic for the theory. The trouble arises because classical OT prohibits intermediate derivational stages – the principle of **strong parallelism**. Strong parallelism requires a rejection of the cyclic interaction between phonology and morphology proposed in many pre-OT generative theories to handle stress preservation, as, under the cycle, stress would be assigned to embedded words like *oríginal*, then to their embedding words while taking the stress pattern of the embedded word into account – *oríginal* is an intermediate derivational stage in the derivation of *orìginálity* (Benua, 1997: 4). Unfortunately for classical OT, it seems to be precisely the cyclic, step-wise process that is vital if stress preservation is to occur – an embedding word cannot logically preserve stress until the stress to be preserved has first been assigned to the embedded word, i.e. preservation cannot happen until there is stress to be preserved! It is clearly very important that, before proposing that OT must be amended in some way to handle English stress preservation, we are sure that preservation indeed occurs; this thesis does just this.

Having shown that there is evidence for English weak stress preservation, I propose that Stratal Optimality Theory ('Stratal OT') (Bermúdez-Otero, 2003, 2007a, b, in preparation; Kiparsky, 1998a, 2000, 2003a, b, 2007b, forthcoming) is the best model of OT in which to handle English stress preservation. Like the theory of Lexical Phonology and Morphology, which received much attention in generative theory during the 1970s and '80s, Stratal OT proposes that there are several ordered phonological and morphosyntactic strata, rather than just a single one. Stratal OT therefore rejects the strong parallelism of classical OT, instead accepting, to a limited extent, the cycle's step-wise derivation. It is argued that the introduction of strata into OT is necessary because attempts to deal with stress preservation without them, notably Benua's (1997) Transderivational Correspondence Theory, are stipulative and limited in their application as compared to Stratal OT.

The case for Stratal OT being made, I argue for a particular version of Stratal OT, proposed by Bermúdez-Otero & McMahon (2006) and Bermúdez-Otero (in preparation). The main reason for arguing this model is that it can handle a new observation made during the empirical investigation of weak stress preservation, which is that weak preservation is a probabilistic phenomenon that is dependent upon word frequency. In the version of Stratal OT advocated here, a functionalist factor like word frequency is able to play a role in the success of weak stress preservation thanks to a new proposal – **fake cyclicity**.

Fake cyclicity is a proposal relevant to specifically the first stratum of the model of Stratal OT, the home of weak stress preservation. In models of Lexical Phonology and Morphology (e.g. Kiparsky, 1982; Giegerich, 1999), this stratum was argued to be internally cyclic in order to handle, among other things, weak stress preservation; as will be seen more than once in this thesis, this proposal is not without its problems. In contrast, fake cyclicity means just what it says: while phenomena like weak stress preservation give the impression of stratum-internal cyclicity on the first stratum, this impression is false, as the stratum is really internally noncyclic. Under fake cyclicity, preserved stress which appears to be the result of cyclic application is instead the result of a combination of redundant lexical storage, redundancy rules and blocking. Fake cyclicity's resemblance to the dual-route model of lexical access and its reinforcement of weak preservation through blocking – blocking being an independently established psycholinguistic phenomenon that is well-known to be affected by word frequency – allow fake cyclicity to capture the demonstrated probabilistic dependence of weak stress preservation upon word frequency.

That the success of weak stress preservation should be partially dependent upon word frequency effects is expected. Stress preservation is an indicator of a relationship between a word and another word embedded within it, and psycholinguistic work shows that word frequency affects how readily words are perceived within others. It follows from gradient morphological complexity that a phonological indication of this complexity – stress preservation – should be variable in its success. Putative pre-tonic stress preservation ((1c), above) has been examined to some extent with respect to word frequency by both Hammond (2003a, b, 2004) and Kraska-Szlenk (2007). The empirical investigation carried out in this thesis is the first analysis of left-edge weak preservation, (1a, b), with respect to word frequency, and the most in-depth analysis of putative weak stress preservation with respect to word frequency to date.

In sum, then, this thesis therefore has two principal outcomes. First, it gives rigorous evidence showing the extent to which weak stress preservation exists in English, and tells us a little more about its precise nature. Second, it supports Bermúdez-Otero's particular version of an emergent theoretical model, Stratal OT.

0.2 Thesis structure

The thesis is structured as follows.

In chapter 1, the theoretical context of the thesis is outlined. The cycle, the key mechanism for handling English stress preservation in earlier Generative Phonology, is outlined in §1.1. Then, in §1.2, the stratal models of Generative Phonology which are the ancestors of Stratal OT are described, along with some of their flaws. A general introduction to Optimality Theory is given in §1.3. The hybrid of stratal models and Optimality Theory, Stratal OT, is then introduced in §1.4, particular attention being given to Bermúdez-Otero's (in preparation) and Bermúdez-Otero & McMahon's (2006) model of Stratal OT. The notion of fake cyclicity is outlined in this section, in §1.4.3. In §1.5, I outline how some of the assumptions made about stress in the thesis relate to previous work.

In chapter 2, the first type of putative weak stress preservation is dealt with – pre-tonic stress preservation. In this chapter, it is argued that there is no evidence for pre-tonic stress preservation, and that we should, instead, think in terms of pre-tonic vowel quality preservation, following Burzio (1994, 2002, 2007). In §2.1, the traditional argument for pre-tonic stress preservation is outlined, before two arguments against pre-tonic stress preservation are introduced in §2.2 and §2.3. The first argument, presented in §2.2, denies the existence of any form of pre-tonic preservation whatsoever; this argument is rejected. However, the second argument, presented in §2.3, proposes pre-tonic vowel quality, not stress, preservation, and this argument is defended, with additions being made to Burzio's arguments for pre-tonic vowel quality, not stress, preservation. In §2.4, some rare instances of what is argued to be genuine pre-tonic stress preservation are presented – forms like *el**è**ctrícity* (*el**é**ctric*). Finally, in §2.5, Hammond's (2003a, b) arguments for pre-tonic preservation being probabilistically predicted by word frequency are presented, as these observations hold whether it is preservation of stress, or of vowel quality directly, in pre-tonic position. Hammond's proposal for phonological identity effects conditioned by word frequency is highly relevant to later chapters of the thesis, where it is argued that another phonological identity effect, left-edge stress preservation, is similarly conditioned by word frequency. The chapter concludes in §2.6.

The second type of weak stress preservation, left-edge stress preservation, is introduced in chapter 3. There is no question of vowel quality preservation here, but there is shown to be controversy over whether and how subsidiary stress assignment at the left edges of words is morphologically conditioned. In §3.1, the stress patterns of words which are not candidates for left-edge stress preservation are outlined, for the purposes of comparison to words where preservation is argued to occur. Then, in §3.2, the first sub-type of left-edge stress preservation – relative prominence preservation – is introduced, and as-yet unresolved arguments against the existence of this type of preservation are presented and queried. The second sub-type of left-edge stress preservation, foot-head preservation, is introduced in §3.3. Potential arguments against this type of preservation are also given, and, again, it is argued that further empirical investigation is required. In §3.4, the particular syllable-weight contexts in which we are likely to discern evidence for left-edge stress preservation are laid out, in preparation for the data analysis which takes place in the following chapters. The chapter is concluded in §3.5.

The results of the first empirical investigation into left-edge stress preservation are presented in chapter 4. In this chapter, words with heavy initial syllables are dealt with, some of which are candidates for relative prominence as well as foot-head preservation. The methods used to collect and categorise the data presented both in this chapter and chapter 5 are presented in §4.1. (The data themselves, taken from pronouncing dictionaries, are included on the appendices at the end of the thesis.) In §4.2, data for heavy-initial words where preservation is expected on the second syllable are presented. The stress behaviour of monomorphemic and bound-root base words is then presented in §4.3; this behaviour is compared to that of the words from §4.2 in order to discern evidence for second-syllable preservation. Heavy-initial words which are candidates for initial-syllable preservation are presented in §4.4, and are analysed for evidence of initial-syllable preservation. The chapter concludes in §4.5.

Chapter 5 presents the results for the investigation into left-edge stress preservation in light-initial words, and takes a very similar format to chapter 4. Data for words where second-syllable preservation is expected are presented in §5.1, and for monomorphemic and bound-root base words in §5.2, before evidence for second-syllable preservation is presented in §5.3. Light-initial words which are candidates for initial-syllable preservation are presented in §5.4. Other disparate types of left-edge preservation are examined in §5.5, before the chapter concludes in §5.6.

Chapter 6 is a data analysis chapter based on the observations from the two preceding chapters; here, the effects of word frequency upon the success of left-edge stress preservation are examined. A frequency analysis approximately comparable to Hammond's (2003a, b) for pre-tonic preservation is presented in §6.1; the outcome for the effect of the frequency of the embedded word conflicts with Hammond's observations, and a possible reason for this is given. In §6.2, precise arguments are made about the relationship between word frequency, lexical access and preservation which can account for the relationships between word frequency and preservation noted in §6.1. Also in this section, Hay's (2001, 2003) hypotheses about relative frequency are introduced; these are explored in relation to left-edge stress preservation in the rest of the chapter. The first relative frequency analysis is presented in §6.3, then further analyses, taking into account the effects of individual suffixes, are presented in §6.4. The plausibility of the results presented in this chapter is evaluated in §6.5, before the chapter concludes in §6.6.

Chapters 7-9 of the thesis deal with the theoretical implications of the preceding empirical investigations and analyses, and of the existence of English stress preservation in general. In chapter 7, Stratal OT is defended as being the best model of Generative Phonology in which to handle English stress preservation. Some of classical OT's strengths with respect to the handling of stress, including stress preservation, are presented in §7.1, before the problems which phonologically opaque phenomena like stress preservation pose for OT are outlined in §7.2. A particularly extensive attempt to handle English stress preservation without strata in OT, Benua's (1997) Transderivational Correspondence Theory, is presented and argued against in §7.3, before Stratal OT is presented and defended against Benua's criticisms in §7.4. The chapter concludes in §7.5.

The particular handling of weak stress preservation in Bermúdez-Otero's model of Stratal OT is presented in chapter 8. In §8.1, Bermúdez-Otero's handling of weak stress preservation using fake cyclicity is given, and I add my own arguments in defence of this proposal, in particular, showing how it can account for the frequency-sensitive nature of left-edge stress preservation that was demonstrated in chapter 6. In §8.2, it is shown how the reinforcement of stress preservation on stratum one can be conceived of as a form of blocking, which is, again, a development of an original proposal made in work by Bermúdez-Otero, as well as being heavily indebted to earlier work in the Lexical Phonology and Morphology tradition, e.g. Giegerich (2001). In §8.3, it is argued that the phonological cycle, the mechanism previously proposed to handle weak stress preservation, cannot cope with the probabilistic nature of weak stress preservation that is demonstrated pretty conclusively in this thesis. Finally, in §8.4, it is shown that fake cyclicity can cope with another phenomenon previously treated as a showpiece case for stratum one's internal cyclicity – Trisyllabic Shortening – before the chapter concludes in §8.5.

The OT constraints which handle weak stress preservation on stratum one, and their manner of operation, are presented in chapter 9. Left-edge foot-head preservation is dealt with first in §9.1, before the more complex case of relative prominence preservation, which requires the introduction of both a revised and a novel constraint, is dealt with in §9.2. In §9.3, it is shown how the failure of left-edge stress preservation comes about; failure is argued to result from both variation in the inputs to the phonology, and as a result of variation in the phonology itself. The problems of coming up with an empirically correct overall ranking of stratum one metrical markedness constraints is explored in §9.4, before the chapter concludes in §9.5.

The thesis is concluded in chapter 10. A brief chapter summary is given in §10.1, before the findings of the thesis are summarised and evaluated in §10.2. Finally, in §10.3, I make some speculative proposals concerning future research.

Part I: previous work

Chapter one: theoretical context of the thesis

1.0 Introduction

This thesis draws upon work in reasonably diverse areas of linguistics, ranging from Metrical Phonology to psycholinguistic work upon word frequency. This chapter does not offer a comprehensive introduction to all of these often disparate areas, as this would not be helpful; instead, much literature is reviewed throughout the thesis as it becomes relevant. The purpose of this chapter is to review the fundamental theoretical contexts of English weak stress preservation and Stratal OT. It begins with a discussion of the phonological cycle, a mechanism proposed by SPE to handle English stress preservation (§1.1), and later incarnations of this mechanism in stratal models of morphophonology (§1.2). Optimality Theory ('OT') is then introduced (§1.3), before the synthesis of strata and OT which is advocated in this thesis – Stratal OT – is introduced (§1.4). Finally, I introduce the assumptions I make about word stress, and show how these relate to previous theory (§1.5).

1.1 The cycle

The cycle has its roots in the founding work of Generative Phonology: Chomsky & Halle (1968) ('SPE').[1] In SPE, the cycle is a principle of phonological rule application (SPE: 15). In a morphologically or syntactically complex form, a sequence of rules will apply to the innermost phonological string that contains no internal morphosyntactic bracketing ([ ]). The innermost set of brackets will be deleted, and the phonological rules will then reapply, again to a string containing no internal brackets. This recursive application of rules continues until the outermost morphosyntactic constituent of the complex form is reached (SPE: chapter 2, especially pp. 15-20; Cole, 1995: 71). So, for a hypothetical string [X[Y]Z], rules would apply first to [Y], then a special rule would delete the brackets immediately either side of Y, before the same rules applied to [XYZ] (Cole, 1995: 71).

Stress has been the primary motivation for the cycle (see Cole's (1995) overview, and references cited therein). In particular, English stress preservation is a

[1] SPE is famed for the cycle, but the principle of cyclic rule application dates to even earlier – Chomsky, Halle & Lukoff (1956) (SPE: 15, f.n. 2).

showpiece cyclic phenomenon (Cole, 1995: §6.1). Given its sensitivity to morphosyntactic structure, the cycle is able to account for the divergence in stress patterns between monomorphemic and morphologically complex words, such as that apparent between *orìginálity* and ***à**bracadábra*. In the cyclic model, the word *originality* has two internal cycles: that which gives *órigin*, and that which gives *oríginal*. The stress assigned on the latter internal cycle to *oríginal* is visible in *orìginálity*. *Abracadabra* is monomorphemic and hence has no internal cycles, thus accounting for the different placements of secondary stress in *orìginálity* and *àbracadábra*. The cycle was also argued by SPE to operate in phrases and compounds, although this latter proposal has since been rejected in Metrical Phonology (e.g. Liberman & Prince, 1977; Giegerich, 1985). (Metrical Phonology has argued that the prosodic structure encoding the stress contours of phrases and compounds can be projected directly from their syntactic trees, removing the need for cyclic rule application – see Liberman & Prince (1977: 258, 332).)

A particular characteristic of the cycle will become particularly controversial in the latter chapters of this thesis. The cycle predicts that, in the environments where stress preservation is permitted, stress preservation will occur whenever it is predicted by a word's morphological structure. In some phonological environments, preservation is consistently prohibited: for example, in English, where degenerate feet would be created by preserving stress on a light syllable immediately preceding another stressed syllable – ***ó**rigin* → ***o**ríginal* (****ò**ríginal*), ***á**tom* → ***a**tómic* (****à**tómic*) (although see Pater, 2000: 268). But, in the phonological environments where preservation can occur – the left-edge and pre-tonic preservation contexts considered in this thesis – the cycle predicts that stress preservation should never fail where it is predicted by a word's morphological structure. This is because the operation of the cycle ensures that a word's stress pattern absolutely reflects its morphological structure, and, as far as the cycle is concerned, a word is either complex, or it is not. It would, therefore, pose a very serious problem indeed for the phonological cycle if there is evidence of stress preservation failure in a phonological context where preservation is permitted, e.g.: *oríginal* → ****ò**riginálity*, cp. *orìginálity* (left-edge preservation); *infórm* → **ìn**for**mátion*, cp. *ìnf**ŏr**mátion* (pre-tonic preservation). We will see in this thesis that such examples of failure do, in fact, occur, and that, consequently, the cycle must be rejected.

Notably, this thesis is by no means the first attempt to reject the phonological cycle. In §1.3, we will see that the classical version of OT is completely opposed to the cycle in any form. Earlier work, which makes a less abrupt break from SPE's rules and derivational depth, has also questioned the cycle's validity. Of particular interest in this respect is the proposal made by Selkirk (1980), as it somewhat resembles the mechanism of fake cyclicity employed later in the thesis. Selkirk argued that predictably complex words were stored in the speaker's memory, along with their full metrical structure. Because, once a word's stress pattern is stored, the mechanism which has generated it – the cycle – is formally redundant, Selkirk rejected the cycle under the principle of Occam's Razor:

> [I]f stress is part of the representation of a lexical item, there is no motivation for a cyclic application of this sort. *Anecdote* has its representation in the lexicon, as does *anecdotal* (Selkirk, 1980: 597).

This is not to say that Selkirk rejected the existence of stress preservation; rather, it was that the mechanism which preserved stresses was historical, and hence no synchronic mechanism like the cycle was necessary:

> The close relation between the stress patterns (i.e. prosodic structure) of the two items would have been established, or laid down, at the point (in time) when *anecdotal* was made from *anecdote*, as a result of word formation (Selkirk, 1980: 598).

As we will see in chapters 6 and 8, Selkirk was right in two ways: first, a word's stress pattern is stored in a speaker's memory; second, this storage does obviate the cycle.[2] However, my interpretation of Selkirk (1980) is that any synchronic relationships between preserving words and their embedded words are rejected altogether. As we will see, fake cyclicity absolutely rejects such an assumption: fake cyclicity proposes that redundant morphophonological relationships

---

[2] For other criticism of Selkirk's proposal, see Hayes (1981). Selkirk's conclusion that the cycle is unnecessary was also reached by Strauss (1983). Strauss proposed that stress assignment is an obligatory morphological adjustment rule, like allomorphy or truncation, rather than a phonological rule. The consequence of Strauss' proposal was that stress preservation occurs, but without the cycle: "the 'cyclic' application of word-internal stress rules in fact is only the outwards appearance of an essentially noncyclic mode of rule application, in which stress assignment applies intrinsically after each Class I suffixation" (Strauss, 1983: 427).

persist between complex words and their bases exist; it is just that, because of the lexical storage of stress patterns, 'cyclic' is not the best description of these redundant relationships. Fake cyclicity will be outlined shortly in §1.4, before being given a full treatment in chapter 8.

1.2 Stratal models

1.2.1 Lexical Phonology and Morphology

1.2.1.1 Overview

Lexical Phonology and Morphology ('LPM') departed from its generative predecessor, SPE, in a very significant way. SPE proposed that syntax constructed both words and sentences, and that the phonology saw all morphosyntactic structure (in the form of the boundary symbols: =, +, #, ##) at once. Following the seminal work of Siegel (1974), LPM proposed that there is a separate morphological component – the lexicon – with a separate block of phonological rules (Chomsky's (1970) Lexicalist Hypothesis).[3] Within the lexicon, morphological and phonological rules are divided in ordered levels or strata. LPM proposed that phonological rules can only see complex morphosyntactic structure if that structure was generated by the current stratum's morphological rules (the 'Bracket Erasure Convention' of Mohanan (1986)), thus rendering the internal morphological structure of words opaque to the phonology of lower strata.[4]

Two key concepts in LPM have been Strict Cyclicity (e.g. Kiparsky, 1982) and Structure Preservation (Kiparsky, 1985: 92). As will be discussed in detail in §1.2.1.2, Strict Cyclicity is a principle of rule application which makes reference to a word's morphological structure: certain phonological rules are only allowed to apply across morpheme boundaries. Structure Preservation also constrains lexical rule application: it proposes that no lexical phonological rule may produce or refer to phonology which cannot be present underlyingly in the language (Kenstowicz, 1994: 221). Under Structure Preservation, Lexical Phonology predicts that allophonic processes must apply postlexically (ibid.).

[3] Siegel (1974) still made use of superfluous SPE boundary symbols; these were replaced by brackets in subsequent models of LPM.
[4] LPM's theory of Bracket Erasure therefore differed from that of SPE, where, in the latter, brackets were erased at the end of each cycle.

In earlier models of LPM, more lexical strata were proposed than have been accepted in later work. Kiparsky (1982) proposed three lexical strata, Halle & Mohanan (1985) and Mohanan (1986) four (see Giegerich, 1999: 2). Later models of LPM – Kiparsky (1985), Booij & Rubach (1987), McMahon (1990), Borowsky (1993), Giegerich (1994, 1999) – have accepted the need for just two.

Most models of LPM have assumed that the morphological side of the stratal organisation of the lexicon is 'affix-driven': affixes are diacritically marked to attach on one or the other of two lexical strata, thus defining the strata. Combined with LPM's serially ordered morphological levels (Siegel's (1974) 'Level Ordering Hypothesis'), the prediction is that no affix marked to attach on a later stratum can attach inside an affix marked to attach on an earlier stratum – the Affix Ordering Generalisation ('AOG') (Selkirk, 1982a). LPM has suffered significantly because of the existence of systematic violations of the AOG (see Aronoff & Sridhar, 1983; Fabb, 1988; Gussmann, 1988; Badecker, 1991; Giegerich, 1999): e.g. occurrence of stratum 1 -*ation* outside stratum 2 -*ise* in *commercialisation*; the necessary attachment of the stratum two prefix *un*- before the stratum one suffix -*ity* in *ungrammaticality* (Hurrell, 2001: 14).[5]

A radical solution to the problems of affix ordering has been proposed by Giegerich (1988, 1994, 1999). Giegerich retains LPM's theory of level ordering, but rejects the affix-driven organisation of the morphology. Instead, Giegerich proposes that morphological strata are defined by morphological base categories like stems and words – the organisation of the lexicon is 'base-driven'.

For English, Giegerich, following Selkirk (1982a), argues that two morphological strata are justified: the Root stratum and the Word stratum. The evidence for this categorical distinction comes from the morphological difference between bound roots like [*matern*-]$_R$ (*maternal*, *maternity*) and free forms like *house*$_N$: affixes like -*ic*, -*ity*, -*ment* can attach readily to bound roots (*sardon+ic*; *amen+ity*; *orna+ment*), whereas more productive affixes like -*ness* and -*less* cannot (a small number of exceptions exist, e.g. *gormless*, *feckless*) (Giegerich, 1999; Hurrell, 2001: 21). All morphological items which undergo stratum one affixation are argued to be roots by Giegerich, so that the category 'root' is recursive and can

[5] [[*Ungrammatical*]*ity*] is an example of a bracketing paradox that goes back to Siegel (1974) (Cole, 1995: 87).

include morphologically complex forms. This is a departure from the traditional definition that may be found in morphology handbooks, where the root is the irreducible, morphologically simple base of a word (see §1.4.2, below, for further discussion). All morphological items which undergo stratum two affixation are words. The category 'word' is similarly (but less controversially) recursive (Giegerich, 1999: 73).

As noted by Hurrell (2001: 18), Giegerich's base-driven model of LPM makes two improvements upon its affix-driven predecessors. First, the base-driven model has no problem with contradictions of the AOG – affixes can have dual-stratum membership. Second, on the phonological side, Giegerich's model allows the operation of cyclic phonological rules to be inherently, rather than stipulatively, constrained. This latter topic is addressed next, in §1.2.1.2.

1.2.1.2 Stratum-internal cyclicity

In the literature on LPM, it is proposed that at least the highest lexical stratum is internally cyclic (examples from models which propose two lexical strata include Booij & Rubach (1987), Borowsky (1993), Giegerich (1994, 1999), McMahon (2000) (Hurrell, 2001: 16, f.n. 3)).[6] This is shown in figure 1:

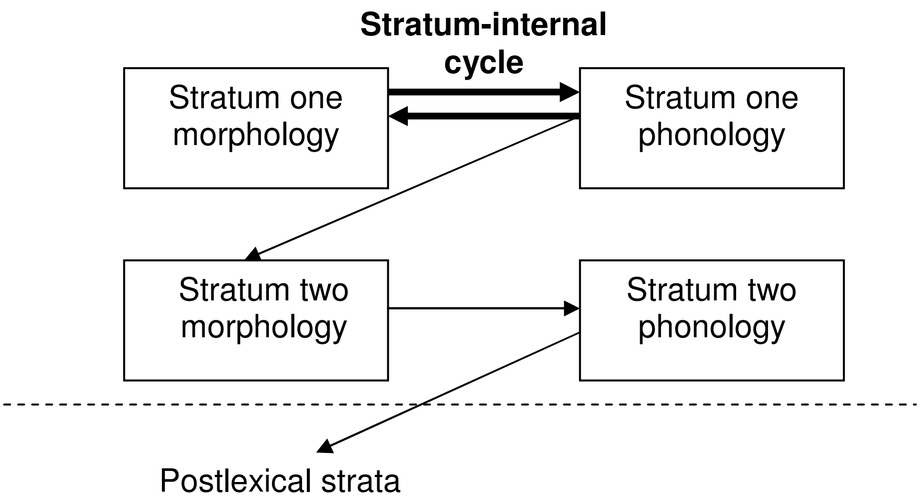


Figure 1: stratum-internal cyclicity

[6] Other models of LPM have assumed that the second stratum is also internally cyclic, e.g. Kiparsky (1982).

Most importantly, the stratum-internal cyclicity of stratum one allows it to handle the weak stress preservation that is the focus of this thesis. The stratum-internal cycle allows complex forms to pass through the stratum one phonology more than once, so that stress can be assigned in a recursive fashion, enabling preservation: stress is assigned first to the innermost constituent, then the next, and so on, until the outermost stratum one morphological constituent is reached (i.e. the form exits onto stratum two). For example, for the complex stratum one form *phenòmenólogy*, stress would be assigned to *phenomenon* on the first pass through the stratum one phonology, giving *phenómenon*. *Phenómenon* would then pass back to the start of the stratum, be suffixed with *-ology*, before passing through the stratum one phonology for the second time. On this second cycle, main stress would be assigned again, but the stress of the embedded constituent *phen**ó**menon* would be preserved, just subordinated – *phen**ò**menólogy*.

Another phonological rule argued to require stratum-internal cyclicity is Trisyllabic Shortening (Kiparsky, 1982). Under Trisyllabic Shortening ('TSS'), vowels are argued to be shortened if they are the stressed head of a trisyllabic sequence in a derived word: e.g. *n***/éː/***tion* → *n***[ǽ]**.*tio.nal*. A crucial argument for the cyclicity of TSS comes from *n*[æ]*tionality*: as the initial syllable of *n*[æ]*tionality* is not the stressed head of a trisyllabic sequence, the only way the initial syllable's short vowel can be accounted for is if it is cyclically inherited from its embedded morphological constituent *n*[ǽ].*tio.nal* (Kiparsky, 1982: 42; Giegerich, 1994: 51, 1999: 101).

One of the key innovations of LPM was to constrain cyclic rule application, thus placing a limit upon derivational abstractness. To this end, the Strict Cycle Condition ('SCC') (Kean, 1974; Mascaró, 1976; Kiparsky, 1982) was introduced. The SCC has been subsequently rejected on empirical grounds (Kiparsky, 1993), and is irrelevant in Stratal OT (§1.4.1); however, the SCC received much attention during the heyday of LPM. The SCC prevented the application of 'structure changing' phonological rules in non-derived environments. TSS was argued to be a structure changing rule (Kiparsky, 1982), and so, under the SCC, could not apply in non-derived forms. This was argued to account for the long vowels in the antepenultimate syllables of non-derived *níghtingale* and *Óberon* (cp. derived *serénity*, *nátional*), and also prevented the proposition of long underlying vowels for nonderived forms like *plátinum* (SPE's 'free ride' problem). Unlike TSS, stress assignment, although cyclic, was argued to be structure-building, not changing, at least on the first cycle: stress assignment to a previously unstressed string adds to, rather than changes, the underlying representation of a word (Kiparsky, 1982).

Kiparsky (1982) proposed that the SCC fell out of the Elsewhere Condition:

(1) The Elsewhere Condition ('EC') (Kiparsky, 1982: 137)

Rule A, B in the same component apply disjunctively to a form ∅ iff:
- (i) The structural description of A (the special rule) properly includes the structural description of B (the general rule).
- (ii) The result of applying A to ∅ is distinct from the result of applying B to ∅

In that case, A is applied first, and if it takes effect, then B is not applied.

Kiparsky proposed that every output of the highest stratum, including complex lexical items, was listed, and that listed lexical item constituted an 'identity rule'. These identity rules constituted rule A of the EC, and any structure changing rule would be rule B. TSS would therefore be blocked in *nightingale* because [naɪtɪngeːl] is an identity rule. In contrast, TSS would not be blocked in derived *sanity* because there is no identity rule [[seːn]ɪtɪ] (*s*[eː]*ne* + -*ity*): the output of the stratum, and therefore the identity rule, is [[sæn]ɪtɪ] (Giegerich, 1999: 103; Hurrell, 2001: 20).

Kiparsky's use of identity rules met with criticism. First, Giegerich (1999: 105) points out that the absence of identity rules from stratum two is stipulative in Kiparsky's model of LPM: the morphology of stratum one and stratum two is exactly the same in Kiparsky (1982). Second, there is the conceptual problem of whether a lexical entry really constitutes a 'rule': "rules do (and identity rules by definition do not), derive something from something else" (Giegerich, 1999: 105).

An alternative proposal has been made by Giegerich (1988, 1994, 1999), which aims to derive SCC from the intrinsic structure of the base-driven stratal model. Key to Giegerich's proposal is his introduction of the Root-to-Word rule. This rule applies to roots prior to their exit from stratum one, turning them into words; these forms are then eligible to proceed onto stratum two. The Root-to-Word rule is a structural necessity in Giegerich's model: without it, the inputs to stratum two would still be roots. Giegerich proposes that Root-to-Word replaces Kiparsky's identity rules, and constitutes a genuine rule A for the Elsewhere Condition. SCC is thus intrinsic, rather than stipulated, in Giegerich's stratal model (Giegerich, 1999: 109-10; Hurrell, 2001: 23).[7]

While a highly principled solution to the stipulative SCC of Kiparsky's model of LPM, Giegerich's Root-to-Word rule has proven to be fatal to stratum one cyclic stress assignment. The Root-to-Word rule, (2), assigns lexical category to previously unspecified roots:

(2) Root-to-Word conversion (Giegerich, 1999: 76)

$[\ ]_R \rightarrow [[\ ]_R]_L$ (L = N, V, A)

English primary word stress is generally acknowledged to be sensitive to lexical category (Hayes, 1982), meaning that primary stress cannot be assigned until after Root-to-Word, on an item's exit from stratum one. The implication of this is that there can be no cyclic stress assignment on stratum one: once a lexical item is a word, it cannot pass again through the Root level's phonology, i.e. stress could not be preserved from the word *oríginal*$_A$ in the complex root *originality*$_R$. Weak stress preservation is therefore impossible in Giegerich's model, unless it is shown that main word stress is not sensitive to lexical category.[8] This avenue is not pursued in this thesis, but, as strong evidence for the existence of weak stress preservation is given in chapters 4 and 5, it appears it is the only option for anyone wanting to retain Giegerich's model with its present theory of strata.[9]

1.2.2 Halle & Vergnaud (1987a, b)

Halle & Vergnaud's (1987a, b) model of cyclic phonology is also stratal. Resembling some of the models of LPM discussed in §1.2.1.2, Halle & Vergnaud propose that

[7] As long as morphology is stipulated to precede phonology. This stipulative ordering is unavoidable in any model of LPM (Giegerich, 1999: §4.2.1).

[8] Such is the argument made for English by Hammond (1999a: 193), and also Burzio (1994).This solution, however, would still involve assigning stress to roots. This may be controversial – see §10.2.1.

[9] This problem of Giegerich's model is discussed at length in Collie (2003) and Hurrell (2001).

there are two phonological strata: one for cyclic phonological processes, and one for noncyclic processes. However, as can be seen from figure 2, Halle & Vergnaud's model differs fundamentally from LPM in that it rejects interactionist morphology and phonology: all morphology takes place before all phonology (Halle & Vergnaud, 1987a: 78, 1987b: 53), resembling SPE's conception of the morphophonology interface.[10]

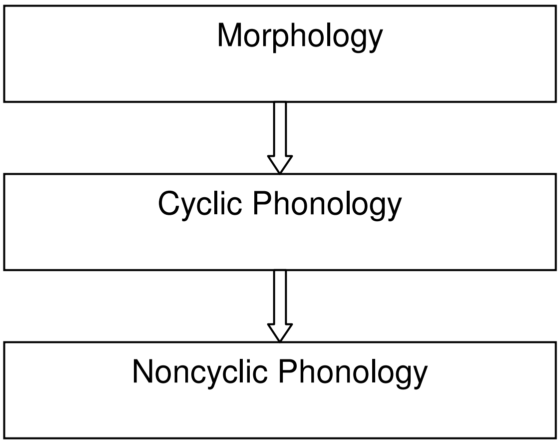


Figure 2: non-interleaved phonology and morphology

Halle & Vergnaud propose that affixes are diacritically marked for whether they trigger cyclic or noncyclic phonological processes. Combined with their model's non-interactionism, this means that Halle & Vergnaud's model rejects the Level Ordering Hypothesis: noncyclic affixation can occur inside cyclic affixation. Violations of the AOG like *ungrammaticality* can therefore be handled in Halle & Vergnaud's model: noncyclic *un-* can attach inside the cyclic suffix *-ity* (Halle & Vergnaud, 1987a: 81).

1.2.2.1 Stratum-internal cyclicity

Halle & Vergnaud propose that their highest phonological stratum is cyclic, resembling the models of LPM discussed in §1.2.1.2. As in LPM, this highest stratum is the home of weak stress preservation.

[10] Halle & Vergnaud's model is sometimes described as a non-interactionist version of LPM. I choose to reserve the term 'LPM' for interactionist models only.

Halle & Vergnaud's theory is built upon Autosegmental Phonology's theory of autosegmental planes. In Autosegmental Phonology (Goldsmith, 1976, 1990; McCarthy, 1986), suprasegmental phenomena like stress and tone each have their own autosegmental plane. All of these different planes intersect, and at this intersection is a string of phonemes, as shown in figure 3:

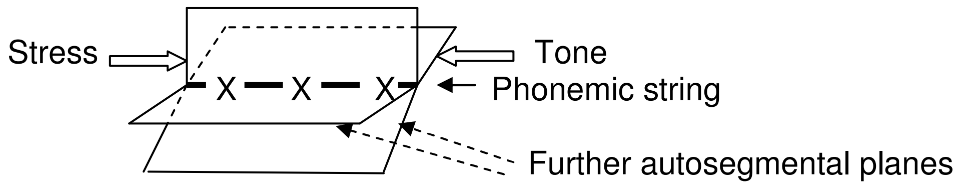


Figure 3: autosegmental planes for a single morpheme[11]

Distinct morphemes are associated with distinct families of planes. A process like affixation will therefore involve the linking of distinct families of planes. Under cyclic affixation, these families of planes remain distinct, as shown in figure 4 (each plane represents the whole family of planes associated with a morpheme that is shown in figure 3):

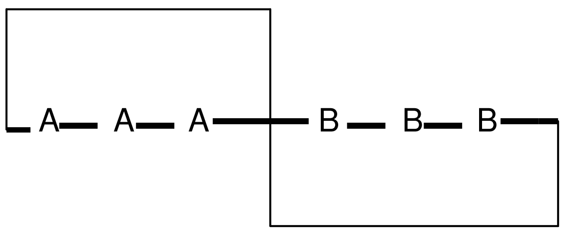


Figure 4: cyclic affixation (I)

Under noncyclic affixation, the affix and base occupy the same plane, as shown in figure 5:

[11] These diagrams are slightly altered from those given in Halle & Vergnaud (1987a: 79, 1987b: 54). Those given here are also those used in Collie (2003).

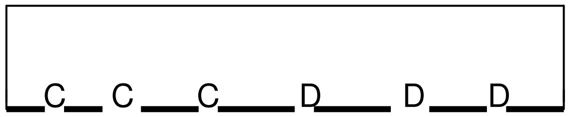


Figure 5: noncyclic affixation

As we can see, the difference between cyclic and non-cyclic morphological processes is that the plane of the base remains visible after cyclic affixation, but is merged with that of the affix under noncyclic affixation.

The distinction between planes is kept under cyclic affixation through the process of plane copy. When a base is affixed with a cyclic suffix, a copy of the base's plane is made, to which the plane of the suffix is added (figure 6):

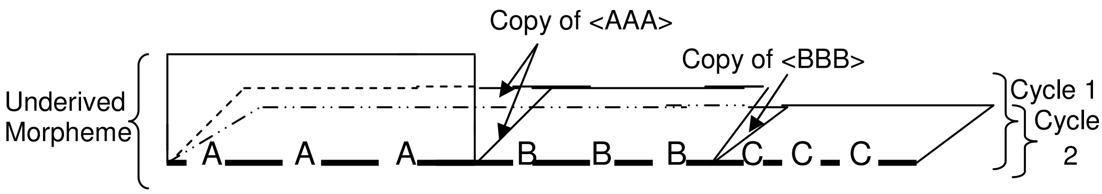


Figure 6: cyclic affixation (II)

Therefore, once all morphological processes are complete, the phonology of the base will still be visible to the phonological module.

Because cyclic suffixation creates new metrical planes, weak stress preservation can occur. For example, with respect to figure 6, <AAA>'s stress plane is still distinct when phonology applies; therefore, the contents of this plane can be reproduced on later replications of the plane via 'Stress Copy':

(3) Stress Copy (Halle & Vergnaud, 1987a: 247)

Place a line 1 asterisk over an element that has stress on any metrical plane.

So, for example, *oríginal*'s stress plane could be copied onto *originality*'s, making preserving *orìginálity* possible. However, Stress Copy is not the default case – the stress of earlier planes is normally deleted under the Stress Erasure Convention:

(4) Stress Erasure Convention (Halle & Vergnaud, 1987a: 83)

> In the input to the rules of cyclic strata information about stress generated on previous passes through the cyclic rules is carried over only if the affixed constituent is itself a domain for the cyclic stress rules [e.g. compound words – SC]. If the affixed constituent is not a domain for the cyclic stress rules, information about stresses assigned on previous passes is erased.

Stress Copy is therefore a "special rule" (Halle & Vergnaud, 1987a: 84; Halle & Kenstowicz, 1991: 490), rather than the default case.

It is clear that the conception of preservation in Halle & Vergnaud (1987a) is altogether different from that proposed in LPM. As we saw in §1.2.1.2, weak stress preservation follows automatically from LPM's cycle internal to stratum one. In contrast, in Halle & Vergnaud's proposal, weak stress preservation is the exception rather than the rule. Halle & Vergnaud defend this approach on the basis of languages like Vedic (1987a: 84-90) (see also their discussion of Lithuanian). In Vedic, there is motivation for two phonological strata from the two classes of suffixes: those which determine main stress ('dominant suffixes'), and those which do not ('recessive suffixes'). Stress-determining suffixes are cyclic, but the accents of the morphemes to which they attach are never present in the surface forms. Halle & Mohanan (1985) have to specify a special rule of accent deletion to cope with this fact; conversely, the desired result comes for free in Halle & Vergnaud's model, where Stress Erasure is the default option.

Halle & Vergnaud's stratal model merits some criticism on account of its handling of cyclicity. Given the null hypothesis of no preservation on the highest stratum, it is unclear whether this stratum is necessarily cyclic at all. For Vedic, for example, Halle & Vergnaud note that "the accentual properties of dominant [cyclic] suffixes other than the rightmost are irrelevant for the location of stress in Vedic words" (1987a: 87). It is therefore unclear why the cycle is required at all for Vedic: given the irrelevance of word-internal morphemes to stress-assignment, stress could surely be assigned just once on the supposedly 'cyclic' stratum, following the addition of the final suffix to a word. As noted by Sainz (1992: 22), this would have damning consequences for Halle & Vergnaud's model, because the only stratal distinction in the model is the diacritic marking between cyclic and noncyclic affixation: if the ‘cyclic’ stratum is, in fact, also noncyclic, the stratal distinction is lost.

The second controversial aspect of Halle & Vergnaud’s handling of cyclicity is Stress Copy. Marvin (2002) points out that Stress Copy seems to be a conceptual backwards step with respect to English, if not other languages.[12] Speaking of the use of Stress Copy in Halle & Kenstowicz (1991) (which assumes the theoretical framework of Halle & Vergnaud (1987a)), Marvin points out that:

> [I]n order to work, the system of Halle & Kenstowicz (1991) has to make use of stress preservation from the earlier cycles, an SPE device that their analysis dispensed with in the first place (Marvin, 2002: 64).

Stress Copy is also contentious in that it allows stress to be copied from ‘**any** metrical plane’, making the device altogether unconstrained (Dresher, 1989: 183; boldface SC). This is particularly controversial with respect to English weak stress preservation: the standard argument is that stress is only preserved from immediately embedded words, e.g. *oríginal* → *orìginálity*, but never from deeper morphological constituents, e.g. *órigin* → **òriginálity* (see especially Benua (1997), discussed in chapter 7 of this thesis). Stress Copy cannot guarantee that English weak stress preservation is restricted in such away. However, it will be argued in chapters 6 and 8 of this thesis that this restriction to preservation from just immediately embedded constituents is, in fact, not necessarily appropriate. In this respect, Halle & Vergnaud’s Stress Copy may be vindicated.

1.2.3 Summary

SPE’s principle of cyclic stress assignment for weak stress preservation has been retained in the stratal models proposed by LPM and Halle & Vergnaud (1987a, b). However, both Halle & Vergnaud’s model and LPM find the handling of weak stress preservation to be problematic. While Kiparsky’s (1982) model of LPM can handle weak stress preservation, it can do so only by recourse to a stipulative version of Strict Cyclicity. As a result of its introduction of a principled theory of Strict Cyclicity, Giegerich’s (1988, 1994, 1999) model of LPM finds itself unable to handle

[12] As we see in chapter 8, Marvin (2002) assumes a version of the cycle and of morphophonology interaction that are both similar in spirit to SPE’s transformational cycle.

weak stress preservation. And, in Halle & Vergnaud's framework, stratum one weak stress preservation is a stipulation upon the stratum's null hypothesis of non-preservation. A stratal theory is, therefore, yet to unproblematically deal with weak stress preservation . This is a problem the present thesis aims to rectify.

1.3 Optimality Theory

The theoretical framework which will be defended in the latter part of this thesis is a particular form of Optimality Theory ('OT') – Stratal OT. Here, the basic concepts of OT are introduced, along with the problem of the classical theory that, in chapter 7, will be argued to make Stratal OT necessary.

1.3.1 Overview

OT (Prince & Smolensky, 2004 [1993]), like SPE and LPM, is a generativist theory. However, OT crucially differs from these earlier theories by capturing phonological generalisations using output constraints, not rewrite rules. OT proposes that languages differ only in terms of their ranking of output constraints; the output constraints themselves are universal.[13]

Many of OT's output constraints are grounded in markedness, where an unmarked value for a particular characteristic is one which is cross-linguistically preferred, and a marked value one which is avoided (Kager, 1999: 2). Markedness constraints make direct statements about the marked or unmarked characteristics of a surface form: e.g. a markedness constraint may be the statement "'syllables are open'" (Kager, 1999: 3). The second major constraint type proposed by OT is faithfulness constraints. Faithfulness constraints try to ensure that outputs of the OT grammar are like their inputs in terms of their phonological characteristics. Faithfulness constraints are necessary to ensure that lexical contrasts present in the underlying representation are present in the output (Benua, 1997: 15; Kager, 1999:

[13] The model of OT outlined here is that presented in the introduction to Kager's (1999) textbook; this model is taken as a classical version of the theory that evolved over the 1990s. There is debate in OT research over most aspects of the OT model presented here, including: constraint grounding; the type and number of output candidates created by Gen; theories of constraint ranking; Lexicon Optimisation; and, last but not least, underlying representations. The model of OT assumed here as 'classical' incorporates the Correspondence Theory of faithfulness (McCarthy & Prince, 1995, 1999) between input and output, rather than the Containment theory of faithfulness (PARSE, FILL) proposed in Prince & Smolensky (2004 [1993]).

5). Finally, there is a third constraint type – alignment constraints. These constraints align the edges of grammatical and prosodic categories, e.g. the edge of a morphological word with a foot.

A central assumption of OT is that constraints are violable: constraints do not embody linguistic universals in the absolute sense. Constraints may make conflicting demands; in a particular grammar, resolution is reached by respecting the demands of the higher-ranked constraint, even if doing so will violate a lower-ranked constraint, and even, with gradiently violable constraints, if this violates the lower-ranked constraint more than the higher constraint ('Strict Domination'). It is the interaction of violable constraints which is absolutely integral to OT (Benua, 1997: 9).

The OT grammar takes the form of an input-output device, shown in figure 7:

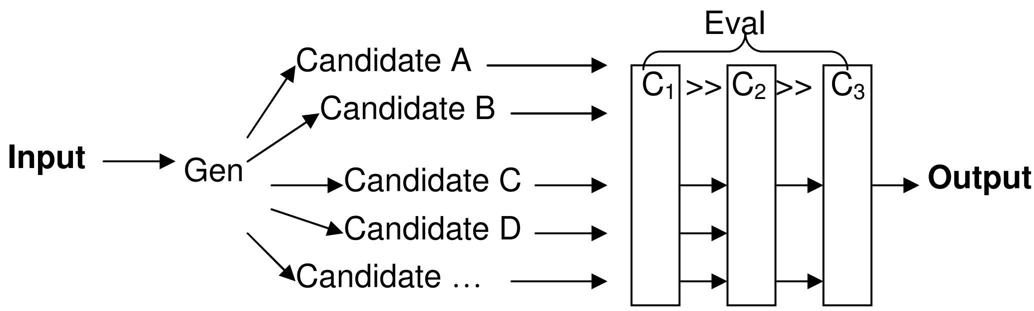


Figure 7: OT grammar (based on Kager, 1999: 9)

The input to the OT grammar, in principle, can be any linguistic form. From the input, a (potentially infinite) group of potential output candidates (A, B, C, etc.) is created by the Generator ('Gen'). There is no consideration of how well-formed these potential outputs are at this stage – Gen is free to generate any linguistic entity ('Freedom of Analysis'). All of the potential output candidates then proceed onto the most important part of the grammar: the Evaluator ('Eval'). Eval consists of a language's constraint hierarchy, plus the way in which output candidates can be evaluated against these constraints.

It is clear from the diagrammatised grammar of figure 7 that all well-formedness statements – i.e. grammatical generalisations – apply at the level of the output only. This is formulated as the principle of 'Richness of the Base' in OT, which ensures that "*all* inputs are possible in all languages" (Prince & Smolensky, 2004 [1993]: 225; italics in original):

(5) Richness of the Base (Kager, 1999: 19)

No constraints hold at the level of underlying forms.

An OT mechanism which is not part of the input-output mapping shown in figure 7, but which is nevertheless vital to the speaker, is the principle of 'Lexicon Optimisation'. Lexicon Optimisation is the means by which a speaker acquires the language-specific underlying representations which form the input to the OT grammar. The version of Lexicon Optimisation proposed by Prince & Smolensky (2004 [1993]: 225-6) states that speakers will choose the input that is most like the output (i.e. incurs the least number of constraint violations), save where there is evidence (e.g. morphophonemic alternations) indicating that an input more disparate from the output should be chosen. Lexicon Optimisation is defined by Prince & Smolensky as in (6):

(6) Lexicon Optimisation (Prince & Smolensky, 2004 [1993]: 225-6)

Suppose that several different inputs $I_1, I_2, \ldots, I_n$ when parsed by a grammar G lead to corresponding outputs $O_1, O_2, \ldots, O_n$, all of which are realized as the same phonetic form Φ - these inputs are all phonetically equivalent with respect to G. Now one of these outputs must be the most harmonic, by virtue of incurring the least significant marks: suppose this optimal one is labelled O_k. Then the learner should choose, as the underlying form for Φ, the input I_k.

Constraint ranking and output evaluation are represented in OT using 'tableaux', an example of which is given in (7) for final voicing in English:

(7) Tableau for English final voicing (from Kager, 1999: 17)

IDENT-IO(voice): correspondent segments in input and output have identical values for [voice]

*VOICED-CODA: obstruents must not be voiced in coda position

Input: /bɛd/	IDENT-IO(voice)	*VOICED-CODA
a. [bɛt]	*!	
☞ b. [bɛd]		*

In a tableau, the constraint ranking is given across the top, with the highest-ranked constraint on the left, and the lowest-ranked on the right. (This is not the case if constraints are vertically separated by dashed rather than solid lines – this indicates the constraints are mutually unranked.) Therefore, for the tableau in (7), IDENT-IO(voice) is highest ranked, then *VOICED-CODA: IDENT-IO(voice) >> *VOICED-CODA. The output candidates put forward by Gen are shown down the left-hand side of the tableau; in (7), two candidates most relevant to the constraints in question, [bɛt] and [bɛd], are shown. Constraint violations are shown by the asterisks, with one asterisk for each violation. A fatal violation – where a candidate is knocked out of the running – is shown by an exclamation mark, and consequently irrelevant violations of lower-ranked constraints are shaded.

The constraints shown in (7) are binary: they are either violated, or they are not. However, many OT practitioners also propose that constraints may be gradient: they may be violated to different degrees. This is especially applicable to alignment constraints; for example, with a constraint like ALL-FT-LEFT (‘align every foot with the left edge of the prosodic word’), a violation will be scored for every syllable by which an individual foot is not aligned with the left edge of the prosodic word. However, the idea of gradient constraints is not uncontroversial – see, for example, McCarthy (2003a).

1.3.2 OT’s problem with English stress preservation

In the model of OT outlined in §1.3.1, just two levels of representation are permitted: Input and Output. This demonstrates the strong emphasis placed upon parallelism in OT since Prince & Smolensky (2004 [1993]): phonological derivations consist of a single mapping from input to output, without any intermediate levels of representation – strong parallelism (Benua, 1997: 11).

OT’s strong emphasis upon strong parallelism is a crucial divergence from earlier, rule-based generative theories. Rules necessarily apply one after the other, making serialism and intermediate representations inevitable. In contrast, in OT's theory of constraint ranking, constraints all express their demands simultaneously, so that "[p]riority among competing goals is modelled as ranking priority of constraints, rather than temporal ordering of rules" (Benua, 1997: 11).[14]

In the context of this thesis, it is most important that, with the rejection of rules, OT has also tended to reject other brands of serialism – strata and the stratum-internal cycle. It is clear that the stratal models outlined in §1.2 could not be interpreted as anything other than serial: each stratum is ordered one after the other, so that intermediate stages unavoidably exist between input and output. Similarly, the stratum-internal cycle is inherently serial: the output of a previous cycle is the input to the next cycle. As noted by both Benua (1997: 11) and Bermúdez-Otero (in preparation), there is no conceptual reason why strata could not be incorporated into OT: it would simply mean that instead of a single, direct input-output mapping, we would have several, one ordered after the other. Nevertheless, the classical model of OT is still monostratal.[15] This move appears to be motivated by cognitive plausibility (Orgun, 1994, 1996a; Bermúdez-Otero, 1999, in preparation): parallel computations have greater cognitive plausibility than serial ones. In this thesis, it is proposed that strong parallelism should be rejected because of the strong evidence presented for English stress preservation. As noted in the introduction, English stress preservation poses a serious problem for a strongly parallel theory of OT: stress preservation appears to be an unavoidably serial phenomenon. The preserving stress pattern of the embedding word (e.g. *orìginálity*) cannot logically be assigned until the stress pattern of its embedded word (*oríginal*) is known – the stress patterns must be assigned in series.

Further evidence for the serialism of English stress preservation is the indication that the preservation relationship is unidirectional: embedding words preserve the stress of their embedded words, but not vice-versa (see especially Benua, 1997: 241). A preservation relationship from embedded to embedding word is necessary to account for the placement of secondary stress in examples like *orìginálity* (*oríginal*) and *ăntìcipátion* (*àntícipate*): in the absence of preservation,

[14] In chapter 7, it is shown that Benua is incorrect to equate serialism with temporal ordering here.
[15] On the definition of the model assumed to be 'classical' here, see f.n. 13.

these words would have initial-syllable secondary stress, as in monomorphemic *àbracadábra* and *Lùxipalílla*. In contrast, no examples have ever been reported where stress preservation from the embedding to embedded word is vital in accounting for the misapplication of stress in the embedded word: it is never the case that stress applies transparently (i.e. as in a phonotactically similar monomorphemic word) in the embedding word, only to misapply in the embedded word, as in the hypothetical example *_**ò**riginálity_ (≈ _àbracadábra_) → *_**ó**riginal_ or *_**ò**rigínal_ (cp. *orígìnal*). While the absence of such examples does not disprove the existence of identity relationships that apply from the derived form to the base, a theory of stress preservation which generates such 'back-copying' relationships should be avoided until evidence for such relationships is found – such a theory would massively over-generate.

The serialism of English stress preservation is therefore fatal to the monostratal model of classical OT presented in §1.3.1. One modified version of OT which has been proposed to handle English stress preservation has been widely adopted within OT: Output-Output Correspondence (see chapter 7). Output-Output Correspondence is an extension of the Correspondence Theory of phonological faithfulness proposed by McCarthy & Prince (1995, 1999). We have already seen Correspondence Theory in operation in the form of faithfulness constraints between input and output (IDENT-IO(nasal) in (7)). Output-Output Correspondence extends phonological faithfulness to relationships between independently occurring outputs of the OT grammar, as shown in figure 8:

OO = Output-Output Correspondence constraint
IO = Input-Output Correspondence constraint

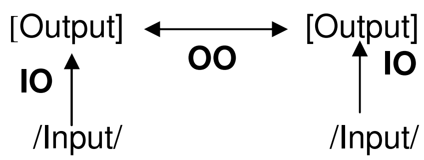


Figure 8: OO-correspondence[16]

[16] In line with Benua (1997: 27), I assume that all outputs, including derived ones, have their own inputs (cp. the diagram shown by Kager (1999: 275)).

Output-Output Correspondence has been argued to handle English stress preservation without compromising OT's strong parallelism. An Output-Output Correspondence constraint relating the independent outputs *original* and *originality* could enforce metrical identity between the two, and therefore stress preservation (figure 9):

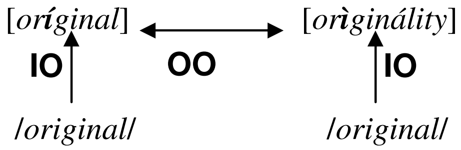


Figure 9: stress preservation and OO-correspondence

It is clear from figure 9 that the Output-Output Correspondence relationship does not appear to be serial in any sense: no intermediate levels of representation are introduced, and the relationship is bidirectional (as shown by the arrows).

The argument for handling English stress preservation using Output-Output Correspondence has been made most extensively by Benua (1997). As we will see in chapter 7, Benua's proposal is extremely problematic where the issues of parallelism and serialism are concerned, and cannot match the restrictiveness and formal economy of the second means of handling English stress preservation in OT: Stratal OT.

1.4 Stratal Optimality Theory

1.4.1 Introduction: OT meets LPM

Stratal Optimality Theory ('Stratal OT') is a hybrid model of OT that broadly combines the insights of stratal models with OT's theory of parallel constraint interaction. In contrast to the theory of classical OT presented in §1.3.1, Stratal OT consists of several serially ordered mappings from input to output (figure 10):

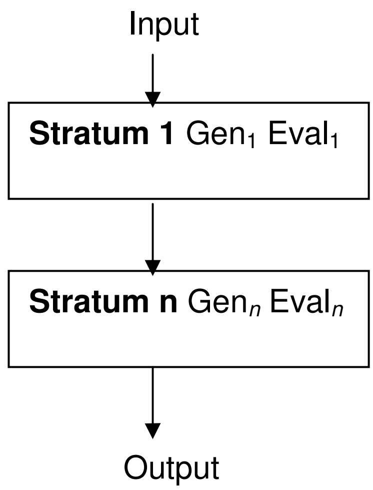


Figure 10: Stratal OT (Kager, 1999: 382)

The particular model of Stratal OT adopted in this thesis is that proposed by Bermúdez-Otero (2003, 2007a, b, in preparation) and Bermúdez-Otero & McMahon (2006). This model argues for the theoretical notion of 'fake cyclicity' (see §1.4.3 and chapter 8), which this thesis argues to be crucial to the handling of weak stress preservation in English. A model of Stratal OT has also been proposed by Kiparsky (1998a, 2000, 2003a, b, 2007b, forthcoming) (termed 'LPM-OT' by Kiparsky). Additionally, Orgun's (1996a, b) Sign-Based Morphology and Phonology may be classed as Stratal OT, although Orgun's model argues for a great many more levels than proposed in either Bermúdez-Otero's or Kiparsky's theories, both of the latter models adopting the restricted number of strata advocated by LPM (Bermúdez-Otero, in preparation) (see §1.4.2). Aside from these three major proposals for Stratal OT, Bermúdez-Otero (1999: 71) notes that combining levels and cycles with OT has been considered by a number of other theorists: McCarthy & Prince (1993a); Black (1994); Potter (1994); Baković (1995); Kenstowicz (1995); Booij (1996, 1997); Lin (1997); Rubach (1997).[17]

Although Stratal OT retains LPM's key concepts of strata and cyclicity,[18] it does not import the theory of LPM wholesale. Being a version of OT, Stratal OT

[17] See also Rubach (2000, 2003, 2004, 2005) (not published at the time of Bermúdez-Otero's writing).
[18] With respect to both strata and cyclicity, it is worth noting that LPM's principle of Bracket Erasure is also retained in Bermúdez-Otero's model of Stratal OT. The difference is that Bracket Erasure is not a stipulation upon the model of Stratal OT, as it was in LPM, but is rather a consequence of independent assumptions about the content of phonological representations and the nonphonological information to which phonology can refer (Bermúdez-Otero, in preparation).

must reject two of LPM's central principles: Structure Preservation and Strict Cyclicity (see §1.2, above). Structure Preservation cannot be retained as an autonomous principle in Stratal OT because it makes crucial reference to constraints upon underlying representations (Kiparsky, 2003a: 256; Bermúdez-Otero & McMahon, 2006: 395):

(8) Structure Preservation (Bermúdez-Otero & McMahon, 2006: (12))
The application of stem-level [stratum one] phonological rules must not violate constraints on underlying representations.

Constraints on underlying representations are not permitted in OT, and therefore Stratal OT, under the principle of Richness of the Base: Richness of the Base requires that constraints hold only at the level of the output.[19] Strict Cyclicity is equally incompatible, because OT cannot restrict phonological processes so that they apply only across morphological boundaries (Łubowicz, 2002);[20] instead, the blocking of phonological processes in nonderived environments must fall out of some ranking of faithfulness and markedness constraints (Burzio, 2000; Łubowicz, 2002). (Bermúdez-Otero (2003) and Bermúdez-Otero & McMahon (2006) also reject Strict Cyclicity on the independent grounds that Kiparsky (1993) shows the principle to be empirically incorrect.)

There is another characteristic of LPM which, while not incompatible with the model of Stratal OT proposed by Bermúdez-Otero, is also not integral to the theory. We saw in §1.2.1 that LPM is an interactionist theory: morphological and phonological levels are interleaved, so that phonology acts upon a morphological constituent immediately after it is created. In contrast, we saw in §1.2.2 that the stratal model proposed by Halle & Vergnaud (1987a, b) retains the non-interactionist principle of SPE: all phonology applies after all morphology. Bermúdez-Otero (in preparation) argues that his model of Stratal OT may be an interactionist or non-interactionist theory (cp. Orgun's (1996a, b) model of Stratal OT, which is absolutely

[19] However, the effects of Structure Preservation are still achieved in Stratal OT: under Lexicon Optimisation, the stem-level constraint ranking controls the content of underlying representations (Kiparsky, 2003a: 256; Bermúdez-Otero & McMahon, 2006: 395). Bermúdez-Otero (2007b) proposes that Chung's Generalisation supersedes Structure Preservation – see §7.4.1.1.
[20] OT also rejects the Elsewhere Condition, from which Strict Cyclicity is derived, as a formal condition on the grammar – see chapter 8.

interactionist). The present thesis takes the same stance. The strongest evidence for interactionism comes from phenomena where a morphological process is sensitive to phonological structure which is not underlying (Siegel, 1974; Aronoff, 1976; Bermúdez-Otero, in preparation); no such evidence is presented in this thesis. Following Bermúdez-Otero, it is assumed in chapter 8 that metrical structure is underlying in stratum one forms. Stress-sensitive suffixation on this stratum (e.g. *arríve+al* but **astónish+al* (Kaisse & Shaw, 1985: 18)) is not a case for interactionism if stress is underlying (Bermúdez-Otero, in preparation).

1.4.2 Number of strata

In Stratal OT, as in LPM, phonological domains correspond to morphosyntactic constructions. In Bermúdez-Otero's model of Stratal OT, as well as that of Kiparsky (1998a, 2000, 2003a, b, forthcoming), phonological domains correspond to the morphosyntactic base categories 'Stem', 'Word' and 'Phrase'. This is a much more restrictive theory of phonological domainhood than that proposed by Orgun (1996a, b), in whose model of Stratal OT "[e]very grammatical construction creates a phonological domain" (Bermúdez-Otero, in preparation) – grammatical and phonological structure is isomorphic. Bermúdez-Otero (ibid.) notes that Orgun's theory of phonological domains is desirable in one respect: phonological domains follow automatically from grammatical structure, without requiring any arbitrary delimitation of which grammatical structures constitute phonological domains, and which do not.[21] But, ultimately, Orgun's theory of domains is undesirably powerful, requiring stipulations to account for situations where the theory would predict phonological domain structure to be too rich (Bermúdez-Otero, in preparation).

A three-level model, with two word formation strata, is in accordance with that ultimately settled upon in LPM (see §1.2.1). Bermúdez-Otero's model also retains one of the most important innovations of Giegerich's model of LPM: morphological levels are defined by morphological base categories, rather than by

[21] It is important to note that the delimitation of phonological domains in Bermúdez-Otero's version of Stratal OT is, nevertheless, principled. Bermúdez-Otero (in preparation) proposes that the delimitation of phonological domains follows the convention of the "Domain Simplification Principle". This condition states that, if a stratum is internally cyclic, every grammatical construction on that stratum constitutes a phonological domain, but if the stratum is noncyclic, it is just the outermost. This condition is similarly seen to apply in models of LPM.

the affixes which can attach on that level. As in Giegerich's model, affixes may have dual stratum membership: the same affix may attach on stratum one and stratum two (Bermúdez-Otero, 2007a: e.n. 9, personal communication).[22]

Bermúdez-Otero's theory of strata does, however, diverge from that of Giegerich (1988, 1994, 1999) in an important way: whereas the first stratum of Bermúdez-Otero's model is the Stem level, that of Giegerich's model of LPM is the Root level. In Giegerich's model (following Selkirk, 1982a), 'root' is a recursive morphological category – roots can be themselves morphologically complex. Bermúdez-Otero's model maintains the traditional notion of 'root' that will be familiar from morphological handbooks and grammars (e.g. Bauer, 1988: 11; Spencer, 1991: 5; Katamba, 1993: 41-6): the irreducible base of a word, such as *desire* in *undesirability*, or *matern-* in *maternity*.[23]

Giegerich rejects a Stem level for English on the grounds that there is no evidence for stems in the language. The morphological category 'stem', in both traditional morphological definitions and that assumed by Giegerich, is a lexical item which is marked for lexical category, unlike a root, but is uninflected, unlike a word (Giegerich, 1999: 89). In a language such as German, there is overt evidence for stems: there exist forms like $[\textit{trink}]_V$ and $[\textit{les}]_V$ which are marked for lexical category, unlike roots, but which are clearly bound because they never occur without overt inflection of some sort (ibid.). Giegerich argues that there is no evidence for bound stems in English:

> As is well known, the distinction between Stem and Word as morphological categories [...] has collapsed in the history of English as part of the decline of the inflectional system (Kastovsky 1992; 1996; Dalton-Puffer 1996; more generally Wurzel 1984): the regular inflection of Present-day English is entirely word-based while the bases of irregular inflection (*cactus*/*cacti* etc.) are adequately analysed as roots (Giegerich, 1999: 88).

---

[22] However, dual-stratum membership of affixes is not seen as the whole solution to violations of the AOG by Bermúdez-Otero (in preparation), *pace* Giegerich (1999). Bermúdez-Otero also rejects LPM's Level Ordering Hypothesis, proposing that Word level construction can occur inside stem-level constructions. Additionally, by incorporating the Prosodic Hierarchy, Bermúdez-Otero's model allows non-isomorphism between morphosyntactic structure and phonological domains, which is known to be yet another solution to violations of the AOG (e.g. *ungrammaticality*) (see Bermúdez-Otero (in preparation) and Hurrell (2001: 40-42), and references cited in both).

[23] Additionally, Giegerich, like others, makes a distinction between bound roots (those which require further affixation before they are eligible to receive lexical category, e.g. *matern-*), and free roots (those which may receive lexical category directly, e.g. *house*). I retain this terminology in this thesis.

Giegerich's argument does not mean that we cannot have a Stem stratum in English; what it does mean is that there is no morphological evidence from English to support such a stratum, so that the stratum would have to be argued for on purely phonological grounds. Such a step means sacrificing the strict congruence between morphologically and phonologically motivated levels that makes a stratal theory like LPM or Stratal OT a highly principled and restrictive one.

In spite of Giegerich's argument, the two morphological strata that will be assumed in this thesis are the Stem level and Word level, as proposed by Bermúdez-Otero (as well most models of LPM), and *pace* Giegerich (1988, 1994, 1999). This is because this thesis presents very strong evidence for the existence of weak stress preservation, and we saw in §1.2.1.2 that, by leaving lexical category assignment until the end of stratum one, Giegerich's Root level is unable to handle English weak stress preservation. Having the highest stratum as the Stem rather than Root level solves this problem: stems, unlike roots, are marked for lexical category, and so cyclic stress could hypothetically occur on the Stem level. Assuming a Stem level is therefore the safest starting position.[24] However, whether Giegerich's model could be salvaged with respect to weak stress preservation in light of the proposal for stratum one 'fake cyclicity' (§1.4.3, chapter 8) is considered in the conclusion of the thesis, in §10.2.1.

In summary, the morphological categories root, stem and word are defined in this thesis as in (9):

(9) Definitions of morphological base categories

Root: A morphologically simple base which is unmarked for lexical category. Roots cannot be inflected.

Stem: A lexical item which is marked for lexical category, but which has not been (phonologically overtly or covertly) inflected. May be morphologically simplex or complex.

[24] It is also worth pointing out that the Stem-and-Word model of Stratal OT proposed by both Bermúdez-Otero and Kiparsky has cross-linguistic applicability: there is evidence for stems in other languages, if not English. For Kiparsky (forthcoming), this cross-linguistic applicability appears to be important: "[i]t goes without saying that this organization [Stem-and-Word] is not specific to Swedish but common to all languages". The debate over whether the number of strata is decided upon language-specific grounds, or is universal, was also an issue in LPM – see Archangeli (1984: 9).

Word: A lexical item marked for lexical category and for inflection.

1.4.3 Fake cyclicity

A particular characteristic of Bermúdez-Otero's model of Stratal OT will be of central importance later in this thesis: the proposal of 'fake cyclicity' for stratum one.

Other stratal models, both of LPM and Stratal OT, assume a particularly rich phonological domain structure for the highest stratum. As we saw in §1.2.1.2, models of LPM have assumed that the highest stratum is internally cyclic; the implication of this is that **every** morphological constituent on this stratum will constitute a phonological domain. This is illustrated in (10) for the stem-level outputs *anticipation* and *originality* ((10) is based on Bermúdez-Otero & McMahon (2006: (14)):

(10) (a) Phonological domain structure if stratum one is internally noncyclic

$[_{\text{stem-level}}$ *anticipation*]

$[_{\text{stem-level}}$ *originality*]

(b) Phonological domain structure if stratum one is internally cyclic

$[_{\text{stem-level}}$ $[_{\text{stem-level}}$ *anticipate*]*ion*]

$[_{\text{stem-level}}$ $[_{\text{stem-level}}$ $[_{\text{stem-level}}$ *origin*]*al*]*ity*]

The stratum-internal cycle is similarly proposed for the Stem level in Kiparsky's version of Stratal OT (e.g. 2003a: 256).

Bermúdez-Otero & McMahon (2006) and Bermúdez-Otero (in preparation) make a radical departure from this tradition of stratum-internal cyclicity: the highest stratum is argued to be internally noncyclic. It is not the case, therefore, that every morphological construction within a word constitutes a phonological domain, as in (10b); it is just the "outermost" (Bermúdez-Otero & McMahon, 2006: 396), as in (10a).

For English, one motivation for the rejection of stratum one's internal cycle is the apparent need for noncyclic phonology on the same stratum. As Bermúdez-Otero & McMahon (2006) note, not all stratum one processes indicate that the stratum-internal cycle is appropriate, e.g. nasal-cluster simplification. Nasal-cluster simplification applies word-finally – *damn* [dæm] – and overapplies before word level affixes – [dæmɪŋ], not *[dæmnɪŋ] – thus confirming its stratum one status.[25] Problematically, nasal-cluster simplification does not overapply before stratum one suffixes: we get [dæmneɪʃn], not *[dæmeɪʃn]. This is not predicted by the stratum-internal cycle, as shown in (11) (from Bermúdez-Otero & McMahon, 2006: (18)):[26]

(11) Incorrect predictions of cyclic nasal-cluster simplification

domain structure	[$_b$[$_a$dæmn]eɪʃn]
inner cycle	dæm
outer cycle	*dəmeɪʃn

However, the rejection of stratum-internal cyclicity need not require the rejection of the weak stress preservation generalisation. The solution lies in Bermúdez-Otero's notion of 'fake cyclicity': the "epiphenomenon of sensitivity to input structure, coupled with the storage of stem-level outputs in the permanent lexicon (nonanalytic listing)" (Bermúdez-Otero, in preparation). A full exposition of fake cyclicity is given in chapter 8, but the proposal will be briefly introduced here.

Fake cyclicity operates as follows. All outputs of stratum one are listed in the permanent lexicon (where, following Bermúdez-Otero (in preparation), 'permanent lexicon' refers to all of the morphological items a speaker memorizes as part of his/her long-term memory). These stored outputs will be referred to as 'lexical entries'. Because these lexical entries are the outputs of the stratum, they will contain all the phonology assigned on stratum one, including stress. This lexical storage

[25] In some models of LPM, it was argued that nasal-cluster simplification was a stratum two process, either applying morpheme- (not word-) finally, e.g. [[*damn*]*ing*] → [[*dam*]*ing*] (e.g. Mohanan, 1986), or by ordering all stratum two phonology before rather than after the stratum two morphology (Borowsky, 1993). Both types of approach have received criticism: see Giegerich (1999: 129-30) and Bermúdez-Otero & McMahon (2006: 395, 398). Giegerich (1999) takes an altogether different approach, and proposes a stratum one process of [n]-insertion that is guided by orthography.

[26] This argument against the stratum-internal cycle is not conclusive – it could be argued that *damnation* is a root-based form, cf. *obfuscation*. Nevertheless, there are still two problems for a stratal model which proposes stratum-internal cyclicity for stratum one: (i) the restriction of stratum-internal cyclicity to the highest stratum; (ii) the probabilistic stress preservation failure explored in the latter chapters of this thesis.

obviates the stratum-internal cycle for stress preservation, as shown in (12) (based on Bermúdez-Otero & McMahon, 2006: (21)):

(12) The stratum-internal cycle versus fake cyclicity

- MAX-FootHead

The output correspondent of an input foot must be a foot head

- ALIGN(ω, L; Σ, L)

Every foot aligns with the left edge of the prosodic word

(a) Stratum-internal cycle

Phonology: 1st cycle[27]

Input: /*phenomenon*/	MAX-FootHead	ALIGN(ω, L; Σ, L)
☞ a. phe(nóme)non		*
b. pheno(ménon)		**!

Phonology: 2nd cycle

Input: /*phe(nóme)non-ology*/	MAX-FootHead	ALIGN(ω, L; Σ, L)
a. (phè.no)me(nó.lo)gy	*!	
☞ b. phe(nò.me)(nó.lo)gy		*

(b) Fake cyclicity analysis

Phonology: only cycle

Input: /*phe(nóme)non-ology*/	MAX-FootHead	ALIGN(ω, L; Σ, L)
a. (phè.no)me(nó.lo)gy	*!	
☞ b. phe(nò.me)(nó.lo)gy		*

In (12b), cp. (12a), the preserving stress of *phenòmenólogy* is derived in a single cycle. *Phe(nómen)on* is listed in the speaker's permanent lexicon as a lexical entry. When the speaker has need to use *phenomenology* for the first time, *-ology* will be added to the lexical entry [*phe(nómen)on*]$_N$. A previous cycle in which the

[27] A third obvious candidate, *(phè.no)(mé.non)*, is not shown. This candidate is assumed to be ruled out by the constraints NONFINALITY ('the final syllable of the word is extrametrical') and FTBIN ('feet are minimally bimoraic'); neither constraint is shown here.

stress of [*phe(nómen)on*]$_N$ is derived from an unmetrified underlying representation, /*phenomenon*/, need not occur – [*phe(nómen)on*]$_N$ is already known.

Remember, however, the starting argument that all stratum one outputs are listed as lexical entries in the speaker's permanent lexicon. This means that [*phe(nòme)(nólo)gy*]$_N$ will itself be listed as a lexical entry. The phonological derivation in (12b) is therefore a 'redundancy rule' (Jackendoff, 1975): the speaker will not perform the derivation in (12b) when he wants to use *phenomenology* in future, but will go straight to the stored stratum one output [*phe(nòme)(nólo)gy*]$_N$. The stratum one constraint hierarchy is therefore anticipated to do very little on-line work. Nevertheless, the persistence of this redundancy rule is vital: it ensures that the weak stress preservation generalisation need not be sacrificed along with the rejection of stratum-internal cyclicity on stratum one.

As we will see in chapter 8, although fake cyclicity is a new proposal, the seeds for it were sown in LPM. The listing of stratum one outputs that is vital to fake cyclicity was also proposed in Kiparsky's (1982), Giegerich's (1988, 1994, 1999) and Borowsky's (1993) models of LPM (see also Mohanan, 1986: §2.6). In particular, the following statement by Borowsky (1993) is prescient with respect to fake cyclicity:

> [A]ll the stems, as well as both classes of affixes, are listed, but no derived Word-level forms occur in the list. [...] The fact that the Stem level is cyclic is due to the fact that the existing lexical items are derived from one another (Borowsky, 1993: 220).

Fake cyclicity's particular contribution to linguistic theory is to fully capitalise upon the suggestion that Stem level forms are stored in a speaker's memory, with the result the stratum-internal cyclicity is rejected. Later in this thesis, it will be shown that weak stress preservation appears to be a probabilistic phenomenon. It is argued that only fake cyclicity, and not the stratum-internal cycle, can capture this observation.

1.5 Stress

In line with work in metrical stress theory (e.g. Liberman, 1975; Liberman & Prince, 1977; Giegerich, 1985; Hayes, 1995), I take stress to be the way in which speech is rhythmically organised: in a string of syllables, some will be more prominent than others, and this variation in prominence provides the rhythm. For example, in the sentence 'This is the house that Jack built', the strong rhythmic beats fall on the first, fourth, sixth and seventh words: ***This*** *is the* ***house*** *that* ***Jack built***. Stress has a number of phonetic correlates, including pitch, duration, and breath pulses; however, there is no single phonetic characteristic which correlates exclusively to stress (Hayes, 1995: 9). Native speakers of a language tend to have pretty good intuitions about the locations of stress in a word (ibid.).

The key phonological unit when dealing with stress is the **foot**; the foot is the next unit in the prosodic hierarchy (Selkirk, 1980) above the syllable. Syllables are organised into feet, with one syllable in each foot being stressed (the 'head'), and the other member syllables being unstressed. In English, feet are left-headed, so that in a polysyllabic foot, the unstressed syllables are to the right of the head: |***This*** *is the* | ***house*** *that*.... A typical foot template in English is the trochee: a left-headed foot consisting of two syllables, (´ x). Recent metrical theory, including work in OT, also recognises a monosyllabic foot consisting of a heavy syllable as a type of trochee (the 'moraic trochee'; Hayes (1995)).

A significant area of debate in metrical theory has been the number of syllables which can be members of a foot. For English, maximally trisyllabic feet have been proposed (e.g. Burzio, 1994); however, most pre-OT analyses propose maximally bisyllabic feet (e.g. Hayes, 1981). With respect to minimal foot size, most pre-OT analyses of English allow monosyllabic feet consisting of a single heavy (but not light) syllable, with the notable exception of Burzio (1994) and Giegerich (1985), both of whom propose that feet are minimally bisyllabic (see also Kager, 1989). With the soft constraints of OT, no parameter on foot size is universally inviolable: for example, the constraint FTBIN, which requires feet to be maximally bisyllabic, could be outranked by a constraint like PARSE-σ, thus forcing trisyllabic or larger feet in a language.

1.5.1 Levels of stress

Work in metrical stress theory proposes that stress is a hierarchical, relational phenomenon – some syllables are more stressed than others. Not only does this manifest as a difference between stressed and unstressed syllables, but also between different degrees of stress. In the thesis, I recognise four different levels of word stress for English: main, secondary, tertiary, and unstressed. By recognising a tertiary level of stress, I am at odds with most recent generative work on stress, and it is my impression that the notion ‘tertiary stress’ is likely to be regarded as obsolete by many current researchers of phonology. I therefore take time to defend my stance here.

The analysis of English stress presented in this thesis is given in OT. OT work differentiates between just three levels of stress: primary stress, main stress, and unstressed (see §9.2.1). This is the stance taken for English by Pater (1995, 2000); Pater argues that this simpler system is necessary because “neither syllable weight, nor stem stress, seem to determine whether a syllable has tertiary or secondary stress” (Pater, 1995: 1, f.n. 1). Similarly, Hammond (1999a) recognises only primary and secondary stress for English, and Burzio (1994), in his constraint-based, non-OT work on English stress, also fails to differentiate between levels of subsidiary stress. In contrast, earlier work in Generative Phonology differentiated between degrees of subsidiary stress for English, as I propose here, for example: SPE; Liberman & Prince (1977); Kiparsky (1979); Hayes (1982); Halle & Vergnaud (1987a); Hammond (1989) (also, more recently, Hayes (1995)).

The differentiation between levels of subsidiary stress is vital to the claims I make about English stress preservation in this thesis. As I discuss in chapter 3, I follow Kiparsky (1979) in arguing that the relative prominences of foot heads may be preserved under morphological embedding and stress subordination: for example, when ***sèn****s****á****tional* (**2**-**1**-0-0) is suffixed with *-ity* to give ***sĕn****s****à****tionálity* (**3**-**2**-0-1-0-0), the stresses of the first and second syllables are both subordinated, but the first syllable remains of a lesser prominence than the second. This situation could not be captured if only one degree of subsidiary stress was recognised, as in ***sèn****s****á****tional* (**2**-**1**-0-0) → ***sèn****s****à****tionálity* (**2**-**2**-0-1-0-0).

In chapters 4 and 6 of the thesis, I give evidence in support of relative prominence preservation. In doing this, I reject Pater’s (1995) argument that preservation (‘stem stress’) does not decide whether a syllable has secondary or tertiary prominence – I show that, at least for relative prominence preservation, it does. In chapter 9, I introduce a way in which the difference in the prominences of subsidiary stresses may be captured in OT.

Chapter two: pre-tonic stress preservation

2.0 Introduction

In this chapter, the status of the first type of weak stress preservation is examined: pre-tonic stress preservation. It is concluded that there is no evidence for pre-tonic stress preservation, and that we must look elsewhere for watertight examples of weak stress preservation.

Pre-tonic stress preservation is argued to occur when the embedded word has main stress which is preserved as subsidiary stress on the pre-tonic syllable in the embedding word: *cond**é**n**se** → cònd**ĕn**sátion*,[1] cp. *cómp**en**sate → còmp**en**sátion*.[2] Since SPE, the inconsistent success of putative pre-tonic stress preservation has been recognised: within a particular phonological context, preservation is argued to be variable, e.g. *cond**én**se → cònd**ĕn**sátion*, but *inf**ór**m → ìn**for**mátion*. In spite of the inconsistent success of putative pre-tonic stress preservation within a phonologically circumscribed context, pre-tonic stress preservation is still recognised in recent phonological theory. Within OT, Pater (1995, 2000) shows that this inconsistency is not an argument against the existence of pre-tonic stress preservation altogether.

Currently more worrying for the status of pre-tonic stress preservation is the argument made by Burzio (1994, 2002). Burzio argues that what is evident in the pre-tonic context in embedding words is not preservation of stress, but rather preservation of vowel quality directly. Under this notion, embedding words never have preserved stress on their second syllables: i.e. *cònd**[ɛ]n**sátion*, not *cònd**[ɛ̆]n**sátion*. Rather, embedding words exhibit preservation of pre-tonic vowel

---

[1] The following symbols are used in this chapter: an acute accent – ' ´ ' – for main stress; a grave accent – ' ` ' – for secondary stress; and ' ˘ ' for tertiary stress. Angular brackets – '< >' – denote extrametricality. Where the level of stress is unimportant, ' ˈ ' is simply used, as in (ˈLL).

[2] SPE (117) also argues that pre-tonic stress preservation occurs in contexts where more than two syllables precede the primary stress and/or it is not the embedded word's primary stress which is preserved: e.g. *ìnstrum**ĕn**tálity ← ìnstrum**én**tal*; *còmplem**ĕn**tárily ← còmplem**én**tary*; *expèrim**ĕn**tátion ← expérim**èn**t*. In this chapter, the focus is upon putative examples of pre-tonic stress preservation which fit the structural description of the Arab Rule (§2.1.1); examples with more than two pre-tonic syllables are given no attention. This is because Pater (2000) gives evidence that examples like *ìnstrum**ĕn**tálity ← ìnstrum**én**tal* and *expèrim**ĕn**tátion ← expérim**èn**t* should not be treated as cases of stress preservation – see f.n. 8, below. However, words with three pre-tonic syllables are used as pre-tonic preservation data by Hammond (2003a) (§2.5.1, below).

quality directly, independent of stress, from embedding words like *cond*[ɛ́]*nse*: *cond*[**ɛ**]*nse* → *cònd*[**ɛ**]*nsátion*.

In this chapter, pre-tonic preservation of vowel quality is argued for over preservation of stress. It is shown that the assumptions which must be made about the nature of vowel reduction if the argument for pre-tonic stress preservation is to be supported cause some serious complications in the analysis of English stress. These complications can be avoided, without introducing different but equivalent disadvantages, if the argument for pre-tonic vowel quality preservation is adopted instead.

The structure of this chapter is as follows. In §2.1, the traditional arguments concerning pre-tonic stress preservation are considered. In §2.2, the arguments which have been made against pre-tonic stress preservation on the grounds of its inconsistency within a phonologically circumscribed context are presented; pre-tonic stress preservation withstands these attacks. However, in §2.3, the case for pre-tonic stress preservation becomes altogether less strong when the argument for direct pre-tonic vowel quality preservation is made. In §2.4, some sparse examples of what is subsequently argued to be genuine pre-tonic stress preservation are presented: cases like *el**é**ctric* → *el**è**ctrícity*. Finally, in §2.5, some recent work on word frequency by Hammond (2003a, b) is presented which attempts to account for why preservation in the pre-tonic context – whether of vowel quality or of stress – is variable within a particular phonological context. Although there are some problems with Hammond's frequency analysis, it does raise the important issue of how probabilistic relationships between words can affect phonological identity between words. The issue of phonological identity and probabilistic relationships between words will be addressed further in chapters 6 and 8.

Unless otherwise indicated, all pronunciations given in this chapter that are taken from pronouncing dictionaries – Jones (2003) or Wells (2000) – are for British English, with British dialectal variants excluded.

2.1 Traditional arguments concerning pre-tonic stress preservation

In this section, three important characteristics of putative pre-tonic stress preservation will be introduced:

(i) The presence of preserved stress on the pre-tonic syllable is inferred by the behaviour of vowel reduction in the pre-tonic syllable, rather than by the presence of salient rhythmic stress (cp. left-edge stress preservation).

(ii) Pre-tonic preservation appears to be phonologically conditioned: the coda quality of the potentially preserving syllable in the embedding word affects whether or not stress is preserved, e.g. *condénse* → *còndĕ**n**sátion* (preservation in a closed, sonorant-coda syllable), versus *defáme* → *dèf**a**mátion* (no preservation in an open syllable).

(iii) Even within any given coda environment, pre-tonic stress preservation is a variable phenomenon: e.g. *condénse* → *còndĕ**n**sátion* (preservation in sonorant-coda syllable) versus *infórm* → *ìnf**or**mátion* (no preservation in sonorant-coda syllable, assuming a rhotic accent).

Point (i) is crucial to the argument that is made against pre-tonic stress preservation later in this chapter. In analyses that support their claim for pre-tonic stress in words like *condensation*, the evidence for pre-tonic stress always comes from the vowel reduction behaviour of the pre-tonic syllable, rather than directly from clearly salient pre-tonic stress. For this reason, SPE holds up putative pre-tonic stress preservation as support for its argument that:

> [S]tress assignment in an early cycle can protect a vowel from phonological reduction, **even when its actual stress**, at the point when the Vowel Reduction Rule applies, **is quite weak** (SPE: 112; see also SPE: 116, f.n. 69; boldface SC).

Similarly, in an extensive treatment of pre-tonic stress preservation, Pater (1995, 2000) infers the presence or absence of pre-tonic stress absolutely on the basis of the presence or absence of vowel reduction or coda coalescence in pre-tonic syllables (e.g. Pater 1995: 1, f.n. 1, 2000: 2). As noted by Burzio (1994: 112), this is not the case with other types of stress preservation – e.g. left-edge stress preservation seen in *orı́ginal* → *orìginálity* – where there is robust rhythmic salience of the preserved stress. Indicating just how questionable the presence of pre-tonic stress is, Kenstowicz (2005) suggests, in passing, that pre-tonic preservation of pre-tonic vowel quality occurs independent of stress:

> In general, the contrast between a full vowel v. schwa is predictable in English as a function of stress; but *comp[ə]nsation* v. *cond[ɛ]nsation* have the same σ̀ σ σ́ σ stress contour and thus raise the question whether English schwa is phonemic after all (Kenstowicz, 2005: 146).[3]

Kenstowicz's proposal for vowel quality preservation independent of pre-tonic stress preservation is exactly that adopted later in this chapter, following Burzio (1994, 2002): in §2.3, we will see that inferring the presence of stress from the absence of vowel reduction alone is not uncontroversial, thus greatly weakening the case for pre-tonic stress preservation.

As an introduction to points (ii) and (iii) – the phonological conditioning and variability of pre-tonic preservation – it is worth considering the original presentation of pre-tonic stress preservation in SPE. SPE (115-117) argues that pre-tonic stress preservation occurs without exception when the putatively preserving syllable is followed by two consonants (SPE: 120) (i.e. where the pre-tonic syllable may be analysed as heavy in metrical theory):

(1) The preservation environment in SPE: pre-tonic heavy syllables

(a) Embedding words which are candidates for preservation

cònd**ĕns**átion ← cond**éns**e   dèp**ŏrt**átion ← dep**órt**

cònd**ŭct**ívity ← cond**úct**   òbj**ĕct**ívity ← obj**éct**ive

cònn**ĕct**ívity ← conn**éct**

(b) Embedding words which are not candidates for preservation

àdj**ect**íval ← ádj**ect**ive   dèm**ons**trátion ← dém**ons**trate

còmp**ens**átion ← cómp**ens**ate

[3] Kenstowicz's proposal of phonemic schwa fits in with Chung's Generalisation that is discussed in chapters 7 and 8 of this thesis. Chung's Generalisation captures the observation that cyclically preserved phonological properties are always those which are lexically contrastive.

Where the embedding words are candidates for preservation – (1a) – it is argued that preservation occurs and that there is stress on the pre-tonic syllable, cp. the words which are not candidates for preservation, (1b).

In contrast, SPE argues that pre-tonic stress preservation appears to routinely fail when the pre-tonic syllable is open (in SPE terms, where the pre-tonic vowel is followed by a single consonant) (SPE: 120-121), as shown in (2):

(2) Routine absence of pre-tonic preservation in open syllables

dèf**a**mátion ← def**áme**	èxpl**a**nátion ← expl**áin**
dìv**i**nátion ← div**íne**	pròv**o**cátion ← prov**óke**

SPE argues that all of the embedding words in (2) have reduced vowels in their pre-tonic syllables, indicating an absence of stress at any level. In this open-syllable phonological context, stress preservation is argued to fail even though the morphological structure of the words in (2) makes them candidates for pre-tonic stress preservation.

In SPE, an apparently problematic situation arises when it is recognised that preservation is not just variable between different phonological contexts, but also within certain particular phonological contexts – the incidence of pre-tonic preservation is no longer predictable. Within the supposedly 'sure thing' pre-tonic heavy syllable environment shown in (1), pre-tonic stress preservation may fail even where predicted by morphological structure, as in *inf**ór**m* → *ìnf**or**mátion*, not *ìnf**ŏr**mátion*.

SPE gets around the *infórm* → *ìnformátion* problem by proposing a different morphological analysis of *information*. *Ìnformátion* is argued to bear no derivational relationship to the verb *infórm*, but rather to be built upon some sort of bound base (SPE: 112, f.n. 64; see also Halle & Vergnaud, 1987a: 251).[4] The justification for this analysis is that:

[4] It is not a central issue in this chapter, but it is worth noting that the unpredictability of pre-tonic preservation is particularly problematic for SPE because SPE enforces morphophonological relationships using the phonological cycle. The cycle predicts that preservation will occur wherever it is predicted by morphological structure, hence SPE's attempt to account for failure in *infórm* →

> [W]e cannot have phrases such as **his information of my friend about the lecture* related to *he informed my friend about the lecture*, as we have *his relaxation of the conditions* related to *he relaxed the conditions* (SPE: 112, f.n. 64).

Similarly, SPE (116, f.n. 69) argues that *condensation* may have both a full and reduced pre-tonic vowel: *cònd*[ɛ]*nsátion* contains embedded *condénse*, and so refers to "'the act of condensing'", whereas *cònd*[ə]*nsátion* refers "'to drops of water on the window pane'", and so is argued not to be the nominalisation of *condénse*. (A similar argument is retained in the cyclic analysis of pre-tonic stress preservation given by Marvin (2002: 68).)

However, as noted by Pater (1995, 2000), not all such instances of pre-tonic preservation failure within a particular phonological context can be explained away on the grounds of alternative morphological analyses:

> Note, however, that *their conservation of the forest*, and *his lamentation of the loss* can be related to *they conserved the forest*, and *he lamented the loss*, even though the pretonic syllables of *lamentation* and *conservation* are reduced (Pater, 1995: f.n. 10; see also Pater, 2000: 261).

Pre-tonic preservation is genuinely and unpredictably variable – morphological structure cannot always predict its incidence.

In summary, the argument for pre-tonic stress preservation is crucially reliant upon the behaviour of pre-tonic vowel reduction. Additionally, from this brief examination of the original SPE proposal for pre-tonic stress preservation, we have observed that pre-tonic preservation is not only conditioned by phonological context, but also varies within a particular phonological context, in spite of predictions made by morphological structure. These latter issues are also crucial to the arguments made against pre-tonic stress preservation later in this chapter, and are presented in more detail next.

2.1.1 Coda-conditioned pre-tonic preservation: the Arab Rule

The Arab Rule (Ross, 1972: 254-258) comes from the original insight by Fidelholtz (1967) about the contrast between two idiolectal pronunciations of *Arab*: [ǽ]*r*[ə]*b*

ìnformátion via an alternative morphological analysis. In chapter 8 of this thesis, the phonological cycle is rejected precisely because of its all-or-nothing nature.

and [éɪ]*r*[æ]*b*. The implication of the Arab Rule is that different vowel reduction behaviour will occur in heavy syllables that immediately follow stressed syllables, based upon two factors:

(i) The weight of the preceding stressed syllable.

(ii) The coda quality of the heavy syllable itself.

The predictions that result from the Arab Rule are summarised in (3) (the examples in (3) are taken from Kager (1989), Pater (1995) and Ross (1972)):

(3) Implications of the Arab Rule

(a) Second syllable vowel reduction occurs when the initial syllable is light, regardless of the coda quality of the second syllable: *Àl*[ə]*xánder*, [ǽ]*r*[ə]*b*, *àr*[ə]*thmétic*, *cáv*[ə]*lcàde*, *clèm*[ə]*ntína*, *Éss*[ə]*x*, *gùar*[ə]*ntée*, *làr*[ə]*ngítis*, *mèl*[ə]*nchólia*.

(b) A full vowel occurs in the second syllable if the preceding syllable is heavy **and** the second syllable has an obstruent coda: e.g. [éɪ]*r***[æ]*b***, *tìckt***[æ]*ck***tóe*, *Tìmb***[ʌ]*c***tóo*; but *ámp***[ə]*r***sànd*, *Àrg***[ə]*n***tína*, *còmp***[ə]*n***sátion*, *Éd***[ə]*n***. Under standard assumptions (e.g. Ross, 1972; Pater, 1995, 2000), a full vowel in the second syllable indicates that it is stressed: e.g. [éɪ]*r*[æ̀]*b, Tìmb*[ʌ̆]*ctóo*, versus *còmp*[ə]*nsátion*.

As can be seen from (3), the Arab Rule is applicable to heavy syllables both word-finally and word-internally. In the context of a preceding light stressed syllable, syllables have a reduced vowel regardless of their coda quality – *Àl*[ə]*xánder*, *Éss*[ə]*x*. With a preceding heavy stressed syllable, syllables have a reduced vowel if their coda is a sonorant, but a full one if the coda consonant is an obstruent – *ámp*[ə]*rsànd*, *Éd*[ə]*n* versus [éɪ]*r*[æ]*b* and *Tìmb*[ʌ]*ctóo*.

The Arab Rule relates to putative pre-tonic stress preservation in a number of important ways. First, it means that if a pre-tonic syllable in a complex word is putatively stressed, it cannot necessarily be argued to be the result of pre-tonic stress preservation: for example, although it could be claimed that the second syllable of *èxp*[ɛ̆]*ctátion* is stressed as a result of preservation from *expéct*, (3b) predicts that this syllable will be stressed on purely phonological grounds anyway, as in monomorphemic *Tìmbŭctóo*. Second, the Arab Rule has the ability to override pre-tonic stress preservation across the board. Both *àdaptátion* and *àffectátion* are reported to have reduced pre-tonic vowels (Pater, 1995: 4), even though they embed the words *adápt* and *afféct*: the phonological effect preceding stressed light syllables, outlined in (3a), overrides preservation. Third, and most interestingly, pre-tonic stress preservation may interfere with the effects of the Arab Rule. As noted by Halle & Vergnaud (1987a: 251) and Pater (1995: 7), the purely phonological requirement for sonorant-coda syllables immediately following a stressed heavy syllable to be unstressed, (3b), may override pre-tonic stress preservation, as shown by the examples in (4) (the examples in (4) are taken from Pater and Halle & Vergnaud):

(4) Pre-tonic sonorant coda syllables in potentially preserving words (rhotic accent)

cònf**ir**mátion	*but*	conf**ír**m	ìnf**or**mátion	*but*	inf**ór**m
còns**er**vátion	*but*	cons**ér**ve	trànsp**or**tátion	*but*	transp**ór**t
còns**ul**tátion	*but*	cons**úl**t	ùs**ur**pátion	*but*	us**úr**p
cònv**er**sátion	*but*	conv**ér**se			

However, pre-tonic stress preservation also seems to have the ability to sometimes override the effect of the Arab Rule, as shown in (5) (the examples in (5) are taken from Kiparsky (1979: f.n. 3), Halle & Vergnaud (1987a) and Pater (1995, 2000)):

(5) The argument for pre-tonic preservation

àdv**ăn**tágeous ← adv**án**tage	àuth**ĕn**tícity ← auth**én**tic
àugm**ĕn**tátion ← augm**én**t	cònd**ĕm**nátion ← cond**ém**n

The Arab Rule predicts that the examples in (5) should have an unstressed vowel in their second syllables; preservation overrides this, resulting in a stressed vowel.

Support for pre-tonic stress preservation therefore crucially occurs in #HH words where the second syllable has a sonorant coda ('#HH(son coda)'). In the phonological context #HH(son coda), there are words like those in (5) which are candidates for pre-tonic preservation, and which have a full vowel in their pre-tonic syllable, contrary to the predictions of the Arab Rule outlined in (3b). In spite of the exceptions to preservation in the #HH(son coda) context like those given in (4), examples like those in (5) appear to be a strong case in favour of pre-tonic preservation.[5] However, there have been challenges to the argument that the words in (5) are evidence for pre-tonic stress preservation. In §2.2, this argument against pre-tonic stress preservation is reviewed and shown to be unfounded.

2.2 Against pre-tonic stress preservation I: no preservation

In §2.1, we saw that the success of pre-tonic preservation varies within the phonological context #HH(son coda). The variable success of pre-tonic preservation within this phonological context helps fuel Halle & Kenstowicz's (1991) argument against pre-tonic stress preservation.

On the basis of the examples given in (6), Halle & Kenstowicz (1991) argue that there is no pre-tonic stress preservation (see also Halle & Vergnaud, 1987a: 233):

(6) Against pre-tonic stress preservation (Halle & Kenstowicz, 1991: 460)

(a) Complex #HH(son coda) words where preservation fails

cònf**ir**mátion, còns**er**vátion, còns**ul**tátion, cònv**er**sátion, ìnf**or**mátion, trànsp**or**tátion, ùs**ur**pátion

[5] Indeed, it may be the rule rather than the exception that complex words like those in (5) display preservation. Halle & Vergnaud (1987a: 251) give three possible vowel quality patterns for *-ation*-derived nouns (i.e. like *augmentation* and *condemnation* from (5)): variably full or reduced; consistently full; consistently reduced. Based on a survey of Kenyon & Knott (1944), they claim that the "largest group, close to the majority" is that with consistently full vowels, i.e. the preserving group.

(b) Monomorphemic/bound-root base words with pre-tonic stress[6]

ìnc**ăn**tátion, òst**ĕn**tátion

What is damning is not the variable failure of stress preservation shown in (6a) (repeated in part from (4)), but the fact that the stress pattern argued to be characteristic of preservation is not exclusive to the preserving context – (6b). Halle & Kenstowicz (1991: 461) conclude that whether a pre-tonic sonorant coda syllable is stressed or not is merely "an idiosyncrasy of individual lexical items", and is not conditioned by morphological complexity in any way.

There are two objections to Halle & Kenstowicz's rejection of pre-tonic preservation. The first is that the examples in (6b) are characterised by "extreme rarity" (Pater, 1995: 20), whereas putative pre-tonic stress in preservation candidates is quite commonplace (see Halle & Vergnaud's (1987a) survey, cited here in footnote 5, and also Liberman & Prince (1977: 299)).[7]

The second argument against Halle & Kenstowicz's rejection of pre-tonic preservation had already been made by Liberman & Prince (1977). Liberman & Prince note that, while we do get preservation failure in the #HH(son coda) context, e.g. *convérse* → *cònversátion*, we do not get the opposite: there is no overgeneration of pre-tonic stresses in complex words, as in *cónc**en**tràte* → **cònc**ĕn**trátion*. As Liberman & Prince (1977: 300-1) point out, this asymmetrical relationship is highly significant: if the pre-tonic stresses in potentially preserving words were really idiosyncratic, and nothing to do with preservation, then *cóncentràte* → **còncĕntrátion* should be just as likely as *convérse* → *cònversátion*. Consequently, to deny that there is pre-tonic preservation would be incorrect.[8]

---

[6] I have removed *ìnc**ăr**nátion* from Halle & Kenstowicz's list, in light of the existence of an apparent base *inc**ár**nate*$_V$ (found in the OED Online, s.v. 'incarnate', not marked as rare or obsolete).

[7] Importantly, as shown in chapters 7 and 8, fake cyclicity predicts lexical exceptions like those in (6b) by Chung's Generalisation. Parallel lexical exceptions are predicted, and do occur, for left-edge stress preservation: *Elízabeth* → *Elìzabéthan* and *Epàminóndas*. In chapters 7 and 8, lexical exceptions like those in (6b) are argued to support a particular analysis of preservation phenomena.

[8] Interestingly, the asymmetry condition does not hold in the potential pre-tonic preservation context with three pre-tonic syllables, e.g. *ìnstruméntal* → *ìnstrumĕntálity*, leading Pater (2000) to argue that a pre-tonic stress in *ìnstrumĕntálity* cannot be attributed to preservation, cp. the bisyllabic pre-tonic context discussed in this chapter. Pater (2000: 266-7) cites *árgum**ent*** → *àrgum**en**tátion* and *éleph**ant*** → *èleph**an**tíasis* as having reduced vowels in the relevant syllable in the embedded word, but a full

In conclusion, the first argument against pre-tonic stress preservation – that pre-tonic stress is not conditioned by morphological complexity – has been defeated. In §2.3, a much more damning argument against pre-tonic stress preservation will be considered: pre-tonic preservation does occur, but it is not stress that is preserved.

2.3 Against pre-tonic stress preservation II: preservation, but not of stress

Like Halle & Kenstowicz (1991), Burzio (1994, 2002) also rejects pre-tonic stress preservation. However, Burzio does so for very different reasons: he argues that there is pre-tonic preservation, but of vowel quality directly, not stress. Under Burzio's argument, the full pre-tonic vowel in *cònd*[ɛ]*nsátion* is not due to the preservation of stress from *cond*[ɛ́]*nse*, with the stress ensuring the unreduced quality of the vowel in the embedding word: *cond*[ɛ́]*nse* → *cònd*[ɛ̆]*nsátion*. Instead, full vowel quality – [ɛ] – is preserved directly: *cond*[ɛ́]*nse* → *cònd*[ɛ]*nsátion*.

For Burzio to be able to reject pre-tonic stress preservation in favour of direct vowel quality preservation, he needs to propose a different relationship between vowel reduction and stress to that which is accepted in the theories that propose pre-tonic stress preservation. Arguments for pre-tonic stress preservation rely upon the standard assumption that there is a symmetrical relationship between vowel quality and stress: full vowels are stressed, and reduced vowels are unstressed. On the basis of nonreduced vowels in pre-tonic syllables, pre-tonic syllables are assumed to be stressed, as outlined in §2.1. Burzio rejects this assumed symmetrical relationship, and, like Fudge (1984) and Hayes (1995: 12), proposes instead that the relationship between vowel reduction and stress is asymmetrical: while a reduced vowel – [ə][9] – is most certainly unstressed, not all unstressed vowels are necessarily reduced. With the argument that not all unstressed vowels are reduced, it can no longer be assumed

---

vowel in the pre-tonic syllable of the embedding word. (I can confirm that [ə] → [ɛ/ə] and [ə] → [ə/æ] – with vowel qualities in order of preference – are given respectively for these examples in Wells (2000) and Jones (2003).) Pater therefore argues that an unreduced vowel in these pre-tonic syllables is not a result of preservation, but rather of "a structural constraint" that also handles monomorphemes like *Kìlimanjáro* versus putative *Hàlicărnássus* (see also Kager, 1989: 123).

[9] SPE also proposes that syllabic sonorants are always unstressed, an argument utilised by Pater (1995, 2000) in his account of pre-tonic stress preservation – see §2.3.2, below.

that the full pre-tonic vowel in *c*ò*nd*[ɛ]*nsátion* indicates the presence of pre-tonic stress.

In §2.3.1, it is shown that the assumption of a symmetrical relationship between vowel quality and stress, necessary to argue for pre-tonic stress preservation, has messy repercussions for the stress system of English. In contrast, as we see in §2.3.2, Burzio's argument for an asymmetrical relationship between vowel reduction and stress can avoid these complications, while providing a well-motivated account of vowel reduction. It will also be seen that Burzio's argument offers greater insight into the coda-quality conditioning of vowel reduction than any argument which assumes that stress alone predicts vowel reduction behaviour.

2.3.1 Problems of assuming a symmetrical relationship between vowel reduction and stress

Following SPE, most major generative treatments of English stress have assumed that there is a symmetrical relationship between stress and vowel reduction (e.g. Ross, 1972; Liberman & Prince, 1977; Hayes, 1981; Halle & Vergnaud, 1987a; Kager, 1989: 140; Pater, 1995, 2000). The assumption that schwa is unstressed is uncontentious (Hayes, 1995: 12). However, analyses which assume a symmetrical relationship between stress and vowel reduction also assume, rather more controversially, that all other vowels – i.e. any vowel except schwa – are always stressed. This assumption has some troublesome implications for English metrical theory.

2.3.1.1 Extrametricality

The first complication of assuming a symmetrical relationship between vowel reduction and stress is to the theory of syllable extrametricality. Hayes (1982: 240) proposes a theory of lexical-category sensitive syllable extrametricality for English, whereby the final syllables of nouns are extrametrical, but those of verbs are not. The advantage of this theory is that, once extrametricality has been used to factor out the difference, the assignment of stress to nouns and verbs has the potential to be identical:

> [W]ith extrametricality, we can capture the unity of stress assignment in nouns on the one hand, and verbs and unsuffixed adjectives on the other (Hayes, 1982: 240).

If a symmetrical relationship between vowel quality and stress is assumed, data cited by Ross (1972: 241-5) is problematic in light of Hayes' theory of lexical-category sensitive syllable extrametricality. Ross gives examples of English nouns where the final syllable contains a vowel (mostly short) followed by a single consonant, and yet the vowel is not schwa – (7). (I have given the range of final vowels cited in Wells (2000) and Jones (2003), in that order, and with the pronunciations in the order of preference given in the dictionaries.)

(7) Full vowels in final closed syllables of nouns (Ross, 1972: 241-3)

albatr**o**ss [ɒ]	Azt**e**c [ɛ]	Beelzeb**u**b [ʌ/ə] [ʌ]	burl**a**p [æ]
cha**o**s [ɒ]	diad**e**m [ɛ/ə]	furl**o**ng [ɒ]	i**a**mb [æ]
Ichab**o**d [ɒ]	Mamaron**e**ck [ɛ]	mayh**e**m [ɛ]	ocel**o**t [ɒ]
Oshk**o**sh [ɒ]	pe**o**n [ə/ɒ]	Tomah**a**wk [ɔː]	

Following the SPE assumption that all vowels except for [ə] must be stressed, Ross (1972) is forced to propose that the final syllables of the nouns in (7) bear stress. Ross proposes that nouns must therefore be marked as to whether they receive final stress, as for the nouns in (7), or not, as in nouns like *Napole*[ə]*n*. Notably, as pointed out by Burzio (1994: 117, 125) (see Schane (2003: 136-7) for a similar argument), the putative final stress in the nouns in (7) has no knock-on effect for the rest of the word's stress pattern: as the reader can easily verify, the nouns in (7) behave exactly as though their final syllable was totally extrametrical (e.g. antepenultimate stress in *Mamároneck* would still follow if this word was metrified *Mamáro<neck>*, rather than *Mamárονèck*).

As Burzio (1994) argues, specifying final stress for nouns like those in (7) breaks down the unification of stress assignment to nouns and verbs. I suggest that if we insist that the full final vowels of the nouns in (7) can only be accounted for by stress, then a rule-based theory will require some sort of late or early stress assignment with respect to extrametricality, or a division between stress assignment processes which see extrametricality and those which do not.[10] In a constraint-based theory like OT, the notion of extrametricality would have to be weakened so that the final syllables of the nouns in (7) are allowed to be feet as long as they are not primary-stress bearing. [11] Whatever the means by which final-syllable stress would be assigned to the nouns in (7), all solutions share the problem that complications to the theory of syllable extrametricality are required purely because of the assumed symmetrical relationship between stress and vowel reduction.[12] The obvious alternative to Ross's argument for final syllable stress for the words in (7) is to argue that it is not stress that is causing the final vowels of these words to be full.

2.3.1.2 Degenerate feet and the bimoraic word minimum

Burzio (1994: 113) notes that the assumption that final syllables with unreduced vowels are stressed can result in degenerate main-stress feet. For example, if the final syllables of *pród*[ʌ/ə]*ct* and *prój*[ɛ/ɪ]*ct* (pronunciations from Wells (2000) and Jones (2003)) are stressed, the main-stressed initial syllable of these words will be a degenerate foot: e.g. ***(pró)****(dùct)*. The problem of word-initial degenerate feet is also pointed out by Pater (2000: 268-9). In the type of example cited by Pater, the second syllable bears main stress, and the initial syllable's vowel is not always schwa. Examples are given in (8). These examples are gathered from several sources which have argued that these words have stressed light initial syllables: Liberman & Prince (1977: 283); Selkirk (1980: 582); Kager (1989: 142); Pater (1995: 35, 2000: 268).[13]

---

[10] With respect to this latter sort of solution, Selkirk (1984: 96-8) argues that the final, extrametrical syllables of nouns like those in (7) receive stress by her Heavy Syllable Basic Beat Rule. This rule, unlike her Main Stress Rule, can assign stress to extrametrical syllables.

[11] The constraint NONFINALITY, often used in OT to capture the skipping of word-final syllables when assigning primary stress, only specifies that the prosodic head of word (i.e. main stress) cannot be on the final syllable (Prince & Smolensky, 2004 [1993]: 48). NONFINALITY does not, therefore, prohibit non-primary stress being assigned to word-final syllables. Nevertheless, the parallels between stress assignment to nouns and verbs would still be lost. An OT version of extrametricality where the word-final syllable is completely unmetrified is proposed by Zamma (2005).

[12] Although there are exceptions to Hayes' noun extrametricality in English, e.g. *Berlín*, *cadét*, *saróng*, these differ from the words in (7) in that the latter require extrametricality to be both visible and invisible to stress assignment; examples like *Berlín* can simply be marked for extrametricality to be consistently inapplicable.

[13] I have removed examples in which the only full vowel would be [ɪ]. This vowel is often in free variation with schwa in English and so could be said to function as a reduced vowel of sorts, but the complication is that it also occurs in stressed syllables in English (Giegerich, 1992: 285).

They include words where a full vowel in the initial syllable could be the result of preservation (8b), e.g. ***á**tom* → ***a**tómic*. Vowel qualities from Wells (2000) and Jones (2003) are given **if words are in either dictionary**; as before, the pronunciation from Jones is given second if it disagrees with that given by Wells:

(8) Light initial pre-tonic syllables with full vowels (sometimes)

(a) Words which are not candidates for preservation

b**a**bóon [æ] [æ/ə] (AmEng)	b**a**ssóon [ə/æ] [ə]	c**a**ffèine [æ] (AmEng)
C**o**lléen [ɒ]	**e**ffáce [ɪ/ɛ/ə] [ɪ/ɛ]	**e**ffête [ɪ/ɛ/ə] [ɪ/ɛ]
erráta [ɛ/ɪ/ə] [ɛ/ɪ]	**E**sséne [ɛ]	f**e**llátio [ɛ/ə/ɪ]
r**a**ccóon [ə/æ]	s**e**ttée [ɛ/ə] [ɛ]	s**u**ttée [ʌ]
t**a**ttóo [æ/ə]	v**a**móose [ə/æ]	

(b) Words which are candidates for initial-syllable preservation

Aarónic	(**Áa**ron)	**a**cídic [ə/æ] [ə]	(**á**cid)
anárchic [æ/ə]	(**á**narchy)	**e**thícian	(**é**thic)
f**a**scístic [æ/ə]	(f**á**scist)	g**e**mmátion	(g**é**mma)
H**e**llénic [ɛ/ɪ/ə]	(H**é**llene)	h**e**ráldic [ə/ɪ/ɛ]	(h**é**rald)
l**e**prótic	(l**é**prosy)	m**a**mmálian [ə/æ]	(m**á**mmal)
m**e**trícian	(m**é**tric)	m**o**dérnity [ɒ/ə/ɪ]	(m**ó**dern)
r**a**bbínic [ə/æ]	(r**á**bbi)		

Pater points out that, if stress were on the light initial syllables of the words in (8), degenerate feet would be created, with serious implications:

> If, however, initial monomoraic syllables can exceptionally be parsed, then there is no reason why a monomoraic syllable that itself makes up a word should not be exceptionally parsed, as it is of course initial. This would be counter to the absoluteness of the bimoraic minimum on words (e.g. */bæ/, */tɛ/ and */pɪ/), which McCarthy & Prince (1986) ascribe to FTBIN (Pater, 2000: 269).

There are ways in which the problem degenerate feet cause for the minimal word requirement could be solved.[14] Within an analysis which assumes a symmetrical relationship between vowel quality and stress, it could be stipulated that degenerate feet are permitted when they are not the only foot in a word. However, this solution would mean sacrificing the generalisation that the bimoraic minimum on English words falls out of the more general prohibition of degenerate feet in English.[15] The obvious alternative to proposing exceptional, word-initial degenerate feet is to argue that these syllables are never heads of feet, even when their vowel is full.

Pater (2000: 269) himself suggests following the latter approach, and rejects that these word-initial syllables with full vowels are the heads of feet. Specifically, Pater proposes that unreduced vowels are stressed **except** when this would require the creation of degenerate feet. This proposed type of nonuniformity can be handled easily enough with OT's theory of violable constraints: as Pater proposes, the requirement for a full vowel to be the head of a foot can take the form of a violable, rather than absolute, constraint. By ranking this violable constraint which requires full vowels to be heads of feet below the constraint FTBIN which prohibits monomoraic feet, it can be ensured that full vowels only correspond to foot heads when it would not result in the creation of a degenerate foot (Pater, 2000: 269). Meanwhile, the continued undominated ranking of FTBIN ensures that content words obey the word minimality requirement. However, clearly, Pater's relaxation of the strictly symmetrical relationship between vowel reduction and stress makes his

---

[14] A means of arguing that the initial syllables of the words in (8) form bimoraic feet is via catalexis. Catalexis is the logical opposite of extrametricality: segmentally empty prosodic constituents are proposed to occur at domain edges (Kiparsky, 1991; Kager, 1995c; re: also Giegerich, 1985; Burzio, 1994). It could be argued that there is a catalectic syllable word-initially for the examples in (8), creating a bimoraic word-initial foot: e.g. *(∅mo)dérnity*. However, word-initial catalexis in a moraic trochaic system, as required to deal with the examples in (8), would result in the catalectic syllable, not the overt word-initial one, being stressed – this is clearly illogical (similar criticisms have been made of Burzio's (1994) word-initial null vowels, e.g. Yamada (1998)).

[15] Although, as noted by Pater (2000: 269), Garrett (1999) proposes, on the basis of an empirical survey, that there is no link between minimal word size and minimal foot size in a language. However, even if Garrett's proposal is accepted, degenerate feet are still undesirable in themselves: it is common for stress systems to avoid or even prohibit degenerate feet (Kager (1995c) and references cited therein; Kager (1995a: 399)). While the violability of OT's FTBIN constraint means that there is no hard absolute prohibition of degenerate feet cross-linguistically in the theory, we should still be wary of introducing degenerate feet into an analysis of English without strong motivation. The only motivation for word-initial degenerate feet in (8) is the assumption that all full vowels must be stressed. Furthermore, in English, the distribution of putative degenerate feet is suspiciously restricted (to just domain-peripheral position) – if they were more widespread, the case for them would be stronger.

proposal for pre-tonic stress preservation entirely stipulative. Once the assumption that all full vowels must be stressed is relaxed, and an asymmetrical relationship between vowel reduction and stress is entertained, it is no more than a theory-internal stipulation to say that stress is present in the pre-tonic context in words like *cond*[ɛ]*nsation*. The absence of pre-tonic vowel reduction can no longer be inferred as evidence for pre-tonic stress preservation. In an analysis which accepts that not all full vowels are stressed, as Pater's now is, such pre-tonic syllables with full vowels could equally be unstressed.

2.3.2 Burzio's (1994, 2002) acoustic-perceptual analysis

Burzio (1994, 2002) avoids stipulations like Pater's by rejecting, from the outset, the assumption that all unreduced vowels are stressed. Burzio proposes that an unstressed syllable may still retain a full vowel if its acoustic energy is needed to aid the perception of an adjacent consonant. Because of this acoustic-perceptual motivation, unreduced vowels may occur in unstressed syllables – stress is not the only factor which determines whether or not a vowel is full.

Burzio's argument crucially relies upon the observation that consonants are reliant upon some sort of vocalic support:

> It can be shown from the fact that (as shown by their citation forms "bee, cee, dee, ef, ...") consonants can in general be articulated only as transitions between openings and closures of the vocal tract, hence in this sense needing vocalic "support" (Burzio, 1994: 114-5).[16]

It may therefore be argued that full vowels preceding consonant clusters – as in the second syllables of *Tim**buct**oo* and *comp**ens**ation* – are necessary to provide vocalic support to their following consonants:

> In closed syllables, corresponding to sequences VC_1C_2, reduction of V would (partially) deprive C_1 of that support, and is for that reason inhibited. In contrast, reduction of V_1 in an open-syllable sequence V_1CV_2 is not comparably inhibited, because support for C is provided here by V_2 (Burzio, 1994: 114-5).

[16] It may be argued that this is not the case with continuant consonants, as these can function as syllable nuclei. This observation has no negative implications for Burzio's acoustic-perceptual argument outlined below: continuant consonants are more sonorous than stop consonants, and so fit with Burzio's predictions.

It is not, therefore, only stress which requires vowels to be full: consonants may also require adjacent vowels to be full in order to aid their own perceptibility, regardless of stress.

Burzio's proposal can account for the distinction in vowel reduction behaviour noted under the Arab Rule. As noted in §2.1.1, under the Arab Rule, obstruent coda second syllables of #HH sequences which are not candidates for pre-tonic preservation tend to have full vowels, e.g. *Timb*[ʌ]*ctoo*. In contrast, those with sonorant codas tend to have reduced vowels, e.g. *comp*[ə]*nsation*, or in some analyses (e.g. Liberman & Prince, 1977; Pater, 1995, 2000) even syllabic sonorants – *còmp*[n̩]*sátion*. In theories which assume a symmetrical relationship between vowel quality and stress, this variation in pre-tonic vowel quality (full versus reduced or absent vowel) is attributed to the presence or absence of stress on the pre-tonic syllable ((3b) above). For example, Pater (1995, 2000) proposes that stress assignment in this pre-tonic context is weight sensitive – only heavy syllables are stressed. The sonorant codas of pre-tonic syllables may coalesce with a syllable nucleus, resulting in a light, unstressable syllable: *còmp*[ɛn]*sátion* → *còmp*[n̩]*sátion*.[17] In contrast, obstruent codas cannot coalesce – *Tìmb*[ʌk]*tóo* → **Tìmb*[k̩]*tóo* – and the resulting heavy pre-tonic syllable receives stress – *Tìmbŭctóo*. Pater's evaluation of this pre-tonic, weight sensitive stress is shown in (9). (Pater does not distinguish between secondary and tertiary stress.)

(9) The Arab Rule as coda coalescence in Pater (1995, 2000)

*OBSNUC = no obstruent nuclei

PARSE-σ = parse syllables

WSP = heavy syllables are stressed

*CLASH-HEAD = no stress on syllables adjacent to the head of the prosodic word

*SONNUC = no sonorant nuclei

[17] Following SPE, Pater (2000: 238, f.n. 1) assumes that syllabic sonorants, as well as schwa, are always unstressed.

Inputs: */compensation/*, */Timbuctoo/*	*OBS NUC	PARSE-σ	WSP	*CLASH -HEAD	*SON NUC
☞ a. (com.p[n̩])⏐sation					*
b (com)(p[ɛn])⏐sation				*!	
☞ a. (Tim)(b[ʌk])⏐too				*	
b. (Tim.b[k̩])⏐too	*!				

In contrast, in Burzio's proposal, the interaction between the vocalic nucleus and the coda consonant alone, **independent of stress**, is enough to account for the presence of a full pre-tonic vowel. The reason sonorant consonants can become syllabic, as is crucial in Pater's analysis, is because of their high sonority value:[18] they require little support from their preceding vowel.[19] In contrast, the low sonority of obstruents means that they crucially require strong vocalic support, and hence tend to be preceded by full rather than reduced (or absent) vowels. A reduced vowel will not provide a following obstruent consonant with sufficient acoustic support, and so would hinder the ease by which the obstruent coda could be perceived (Burzio, 1994: 115). Crucially, an obstruent coda's need for vocalic support means that a preceding full vowel will be preferred over a reduced one **independent** of whether the syllable in question is stressed or not.

Burzio (2002) shows that his more complex account of vowel reduction follows from the principles of Steriade's (1994, 1997) Dispersion Theory. Dispersion Theory proposes that the perceptual distance between sounds must be as great as possible: weak contrast of sounds is avoided.[20] Should the perceptual distance not be great enough, Dispersion Theory proposes two solutions: enlarge the distance, and therefore the contrast; or get rid of the perceptual distance, and therefore the contrast, altogether (Burzio, 2002: 3). Under the principles of Dispersion Theory, vowel reduction is anticipated in unstressed syllables: unstressed syllables have less

[18] On the sonority of sonorant consonants see e.g. Steriade (1982) and Selkirk (1984) (Burzio, 1994: 115).
[19] Whereas Pater's argument for syllabic sonorants seems to imply that there is no remaining vowel preceding the sonorant nucleus, Burzio (2002: 11) proposes that sonorants still require some vocalic support, e.g. schwa. However, as Burzio (2002: 13) proposes that syllables with reduced vowels are lighter than those with full vowels (2002: 13), his argument converges with Pater's observations about pre-tonic syllable weight.
[20] For this original idea see Lindblom (1986).

acoustic energy (see (11) below), and so any vowel contrasts in them will be weak, and neutralisation of the vowel contrast to schwa will be encouraged (Burzio, 2002: 10).[21] However, this vowel reduction will not go unopposed: vowel reduction would decrease the vocalic support for any following obstruent consonant which crucially requires the greater acoustic energy of a full vowel to aid its perceptibility – reduced vowels are less sonorous than full vowels. Under the assumption that obstruent codas are not permitted to undergo neutralisation themselves in English (Burzio, 2002: 10), the vowel is required to stay full to maximise the perceptual contrast of the obstruent coda. (The only exception to this concerns coronal obstruents – an important issue which will be returned to shortly.) Conversely, the greater inherent acoustic energy of sonorants means that they are intrinsically more perceptually salient, and consequently the supporting vowel is more free to neutralise to schwa in an unstressed syllable.

Burzio's acoustic-perceptual account of vowel reduction can also extend to the other Arab Rule context: reduction in the vowel of the second syllable of ˈLH sequences, whatever the quality of the second syllable's coda. Burzio's account relies upon the following two principles:

(10) Syllable weight and acoustic energy (Burzio, 2002: 13)
Syllable weight is commensurate with acoustic energy: heavy syllables with reduced vowels are less heavy than those with full vowels.

(11) ΔE (Burzio, 2002: 4)
Maximise the [acoustic] energy difference between stressed and unstressed syllables.

Principle (10) is an assumption made by Burzio, although it does in fact resemble Gordon's (2002: 52) observation that central, schwa-type vowels may function as

---

[21] This strategy of vowel reduction to overcome the weak perceptual contrast in unstressed syllables is not universal: neutralisation phenomena vary between languages, as shown by Burzio's (2002) comparison of English and Italian. The analysis of English vowel reduction which is presented here is given in terms of violable constraints in Burzio (2002), the ranking of which varies between languages so as to predict language-specific neutralisation strategies.

lighter than peripheral vowels.[22] Sainz (1992: 40-45) also proposes that schwa is lighter than other vowels. Principle (11) is the formalisation of Burzio's (in press) empirical observation (based upon measurements of acoustic energy and analysis of spectrograms) that unstressed syllables have reduced acoustic energy as compared to stressed syllables (see also Gordon, 2002; de Lacy 2006: 225).

On the basis of principles (10) and (11), vowel reduction will be required in the second member of an ˈLH sequence. If, as is proposed in (10), syllable weight is commensurate with acoustic energy, then the light syllable will realise a lower level of acoustic energy than the heavy second syllable. Given, however, that it is the light syllable and not the heavy second syllable which is stressed, we will have a violation of (11): rather than getting the desired reduction in acoustic energy as we move from stressed to unstressed syllable, the reduction will be small, or we may even get an increase. One way to solve this undesirable mismatch between the energy and stress profiles is to decrease the acoustic energy of the second syllable. This can be achieved by reducing the vowel in the second syllable to schwa: schwa has less acoustic energy than a full vowel.

In support of Burzio's argument, and, in particular, the assumption made in (10), it is notable that (ˈLH) trochaic feet appear to be cross-linguistically marked (e.g. Kager, 1989; Prince, 1990; Hayes, 1995; Alber, 1997). If reduction of the acoustic energy of the second syllable of an (ˈLH) foot, via vowel reduction, does decrease the second syllable's weight as proposed by (10), then this would be a move towards the canonical trochee, (ˈLL).

Although Burzio's account of the *Ár*[ə]*b* phenomenon therefore seems reasonable, it is not the case that this phenomenon went unaccounted for in theories which assume a symmetrical relationship between vowel reduction and stress. Pater (1995, 2000) proposes a principled solution to this problem, shown in (12):

[22] Furthermore, Gordon (2002: 52-3) argues that the phonetic **durational** difference between central vowels like schwa and peripheral vowels plays a part in establishing a weight distinction. English is well-recognised as distinguishing between long and short vowels in terms of weight, and so the schwa-peripheral distinction could be argued to fit in with this established weight distinction.

(12) Vowel reduction with #LH in Pater (1995)

Inputs: /*Alexander*/	*OBS NUC	PARSE -σ	WSP	*CLASH-HEAD	*SONNUC
☞ a. (Alex)⏐ander			*		
b. (Alx)⏐ander	*!				
c. A(lex)⏐ander		*!		*	

The second syllable of *Àlexánder* is not stressed, in spite of the second syllable having an obstruent coda – this would leave the word-initial syllable unparsed, as in candidate c. **A(lèx)(án)<der>*. Under the symmetrical relationship between vowel reduction and stress which Pater assumes, the vowel of the second syllable will therefore reduce, as it is predicted to be unstressed.

What I have shown in this section is that Burzio's acoustic-perceptual analysis offers plausible accounts of vowel reduction in Arab Rule contexts, without incurring the problems of an analysis which assumes that all full vowels must be stressed: degenerate feet, and stressing of supposedly extrametrical syllables. Burzio's proposal of an acoustic-perceptual argument over a stress argument is particularly intuitive when we consider that coda-consonant quality has been argued to play a central role in determining whether pre-tonic syllables have reduction-preventing low-level stress. As noted by Burzio (1994: 125), when we are in no doubt that a syllable is stressed – for example, main stress, or secondary stress in *àbracadábra* or *orìginálity* – the quality of the coda consonant plays no part in determining stress assignment. Rather, it seems to be the case that coda-place sensitive stress assignment must be invoked in precisely the contexts where the presence or absence of vowel reduction is argued to be the only indicator of stress placement. This has been the case with pre-tonic syllables so far; we will now see in §2.3.2.1 that it applies to word-final syllables also.

2.3.2.1 Word-final coda-conditioned vowel reduction

It is now shown that Burzio's (2002) acoustic-perceptual account of vowel reduction can explain instances of coda-conditioned reduction which are problematic for a stress account.

Ross (1972: 250-1) notes that vowel reduction in word-final syllables occurs freely preceding the coronal stops /t, d/. Examples from Ross, and also Burzio (2002), are given in (13), with pronunciations verified in Wells (2000) and Jones (2003). (For both (13) and (14), Wells is given first, then Jones, where pronunciations disagree between dictionaries; variable pronunciations are given in the order of preference cited by the particular dictionary.)[23]

(13) Word-final coronal stops

carp**e**t [ə] (U.S. Eng)	chari**o**t [ə]	chevi**o**t [ə]
Connectic**u**t [ə]	idi**o**t [ə]	Ili**a**d [ə/æ] [æ/ə]
Lillip**u**t [ʌ/ə]	Mohamm**e**d [ɪ/ə/ɛ]	myri**a**d [ə]
peri**o**d [ə]	pil**o**t [ə]	po**e**t [ə] (U.S. Eng)

While schwa does not occur across the board in (13), it is very much prevalent.

Ross (1972: 250-2) proposes that vowel reduction is more inhibited preceding the word-final velar and labial stops /p, b, k, g/. Examples from Ross and Burzio are given in (14), again with pronunciations taken from Jones (2003) and Wells (2000).

(14) Word-final velar and labial stops[24]

Ah**a**b [æ]	Azt**e**c [ɛ]	baob**a**b [æ]
beb**o**p [ɒ]	Beelzeb**u**b [ʌ/ə] [ʌ]	Cant**a**b [æ]
Carn**a**p [æ]	demag**o**g [ɒ]	handic**a**p [æ]
humb**u**g [ʌ]	kay**a**k [æ]	kidn**a**p [æ]
ketch**u**p [ə/ʌ] [ʌ/ə]	kop**e**ck [ɛ]	lollyp**o**p [ɒ]
Mamaron**e**ck [ɛ]	musk**e**g [ɛ]	nab**o**b [ɒ]
nutm**e**g [ɛ]	pollyw**o**g [ɒ]	satr**a**p [æ/ə]

---

[23] I have removed examples which only occur with final [ɪ] from both (13) and (14), re: f.n. 13.

[24] Burzio (2002) also notes the existence of word-medial examples, e.g. *aut**o**psy*, *archit**e**ctonic*. In contrast, although the argument for [t, d] versus other stops is similarly not restricted to word-final position, in practice, it seems to be relevant only word-finally: "coronal stops in word-medial codas (e.g. *Watkins*) are rare altogether – an accident, from the present perspective" (Burzio, 2002: 10).

scalaw**a**g [æ] shamr**o**ck [ɒ] wicki**u**p [ʌ]

In (14), it can be seen that, while instances of schwa are present, these are the exception, not the rule, cp. (13).

In an analysis which assumes a symmetrical relationship between vowel quality and stress, any final-syllable full vowels must bear stress. This leads Ross (1972) to propose that certain stress rules are directly sensitive to coda-place to capture the difference in the general patterns of vowel reduction behaviour between (13) and (14).[25] In Burzio (2002), the variation in vowel reduction behaviour seen between (13) and (14) follows naturally from the acoustic-perceptual, dispersion theoretic analysis. Burzio (2002) proposes that the final syllables of the words in both (13) and (14) are unstressed. This lack of stress means that the vowel in the final syllable has lower energy, lessening its perceptibility, and so encouraging its neutralization to [ə] in order to avoid weak contrast. However, the vowel of the final syllable also has another function which conflicts with reduction – to support the following coda consonant. Under Burzio's assumption that, in English, there is pressure to preserve the coda consonant's underlying place of articulation, the preceding vowel will need to be full in order to cue the non-neutralised coda consonant. We can therefore account for full vowels in the final syllables of the words in (14), e.g. *pollyw*[ɒg], *handic*[æp], *Mamáron*[ɛk]: the final obstruents of these words are not sonorous enough to cue themselves, and so require the energy provided by full preceding vowels to cue their non-neutralised place of articulation.

This still leaves the reduced vowels preceding coronal stops, (13), to be accounted for. The case of coronal stops is special. Coronal stops are not sonorous enough to stand alone, so we would expect them to require a cue from a preceding vowel, just like velar and labial stops. However, with coronal stops, there is no need to avoid neutralisation of their underlying place of articulation: the unmarked,

[25] There are a well-defined group of exceptions to (14): *Arab*, *cherub*, *syrup*, etc. As Ross (1972: 252, f.n. 20) notes, vowel reduction is predicted in these syllables under the ˈLH condition of the Arab Rule.

neutralised place **is** coronal (Paradis & Prunet, 1991).[26] Underlyingly coronal stops will not require a full vowel to cue them, as they do not need to be protected from neutralisation – they are already there. As a result, reduced vowels are likely to occur preceding the coronal stops /t, d/, but not preceding the non-coronal stops /p, b, k, g/. The variation in vowel reduction behaviour between (13) and (14) is therefore explained in Burzio's acoustic-perceptual analysis.

The problem with a stress analysis of the vowel reduction behaviour of the word-final syllables in (13) and (14), like that proposed by Ross, is that, based upon what we have seen so far, it appears to be no more than stipulative. Stress assignment follows from suprasegmental distinctions: some form of syllable counting and/or the contrast between heavy versus light syllables. To my knowledge, the only way in which the coda quality of word-final consonants could influence stress assignment to word-final syllables is if coda quality plays a part in determining the weight of these word-final syllables, and the mechanism which assigns word-final stress is weight sensitive. Pater (1995, 2000) shows that the putative sonorant versus obstruent coda contrast in pre-tonic stress assignment (putative *Tìmb**ŭc**tóo* versus *còmp**en**sátion*) can be handled this way (see §2.3.2): differences in pre-tonic stress placement, and so, in Pater's analysis, pre-tonic vowel reduction, may be accounted for by the effect of coda coalescence upon pre-tonic syllable weight. However, there is no way in which a similar analysis could be invoked here – none of the word-final stops in question are subject to coalescence.

In a stress analysis of the word-final syllables in question, an alternative weight distinction has been proposed by Elfner (2007). The argument in moraic theories of faithfulness has been that, within any language, codas are either heavy, or weightless: there seems to be no language where there is contrastive weight of CVC depending upon the quality of the coda consonant (Elfner, 2007: 1). Elfner argues

---

[26] There is, of course, controversy over markedness in recent research. Standard arguments for coronals being unmarked include their patterning differently as compared to other places of articulation (Rice, 1999a, citing Paradis & Prunet, 1991). Rice also cites her own work with Peter Avery on cross-linguistic segmental inventories: "with rare exception, languages include a coronal obstruent, and the presence of other places of articulation implies the presence of coronals" (Rice, 1999b: 3). Central vowels like schwa are a recognised result of the 'emergence of the unmarked' in neutralisation (Rice, 1999a: 5).

that this is not the case for English, proposing, instead, that the data in (13) and (14) is evidence of contrastive coda-consonant weight: non-coronal obstruents always count as moraic, but word-final coronal obstruents may or may not count as moraic. Elfner's analysis is reliant upon non-exhaustive syllabification: consonants which could function as syllable codas may instead be attached directly to the prosodic word (following Selkirk's (1996: 190) proposal for nonexhaustive prosodic domination).

The OT constraints which Elfner uses are given in (15), of which the PARSESEG constraints are new innovations:

(15) Elfner's (2007) constraints

PARSESEG[+cor]: assign one violation mark for every [+ coronal] segment that is not dominated by a syllable node.

PARSESEG[-cor]: assign one violation mark for every [-coronal] segment that is not dominated by a syllable node.

PARSEμ: assign one violation mark for every mora that is not dominated by a syllable node.

WBYP: Coda consonants are moraic.

DEPμ: no moras should be present in the output that should not be present in the input.

MAXμ: no moras should be deleted in the output which are present in the input.

*Cμ: consonants are not moraic.

Under Richness of the Base, final coronals may be underlyingly moraic or non-moraic. Elfner's proposed constraint ranking ensures this variation will be present in the output, as shown in (16) (I use { } brackets to indicate consonants attached directly to the prosodic word level as appendices):

(16) Variable weight of word-final coronals in Elfner (2007: 6)[27]

(a) Underlyingly non-moraic final coronal

Input: /po$_{\mu\mu}$ə$_{\mu}$t/	WbyP	Depμ	ParseSeg[+cor]
☞ a. po$_{\mu\mu}$.ə $_{\mu}${t}			*
b. po$_{\mu\mu}$.ə $_{\mu}$t	*!		
c. po$_{\mu\mu}$.ə $_{\mu}$t$_{\mu}$		*!	

(b) Underlyingly moraic final coronal

Input: /lɪ$_{\mu}$lɪ$_{\mu}$pʌ$_{\mu}$t$_{\mu}$/	Maxμ	*Cμ	Parseμ
☞ a. lɪ$_{\mu}$.lɪ$_{\mu}$.pʌ$_{\mu}$t$_{\mu}$		*	
b. lɪ$_{\mu}$.lɪ$_{\mu}$.pʌ$_{\mu}${t$_{\mu}$}		*	*!
c. lɪ$_{\mu}$.lɪ$_{\mu}$.pʌ$_{\mu}${t}	*!		

By ranking ParseSeg[-cor] above Depμ, rather than below it like ParseSeg[+cor], non-coronal final obstruents are ensured to always be moraic, even if they are underlyingly non-moraic; this is shown in (17):

(17) Consistent weight and syllabification of word-final non-coronals (Elfner, 2007: 7)[28]

Input: /skæ$_{\mu}$lɪ$_{\mu}$wæ$_{\mu}$g/	WbyP	ParseSeg[-cor]	Depμ
☞ a. skæ$_{\mu}$.lɪ$_{\mu}$.wæ$_{\mu}$g$_{\mu}$			*
b. skæ$_{\mu}$.lɪ$_{\mu}$.wæ$_{\mu}${g}		*!	
c. skæ$_{\mu}$.lɪ$_{\mu}$.wæ$_{\mu}$g	*!		

The overall ranking of constraints proposed by Elfner (2007: 7) is given in (18):

[27] Of course, under Richness of the Base, the opposite situation is also possible: an underlyingly moraic final /t/ in *poet*, resulting in a full vowel, and a non-moraic one in *Liliput*, resulting in a reduced vowel. However, Richness of the Base does not require that every conceivable lexical item actually exists; it just requires that no systematic restrictions, such as 'no underlyingly moraic coronal consonants', apply to underlying representations.

[28] Elfner (2007: 7) does not show WbyP or candidate c. in this tableau.

(18) Ranking of constraints proposed by Elfner

WbyP   ParseSeg[-cor]   Maxμ
  \   / |
   Depμ |
    | |
ParseSeg[+cor]   *Cμ   Parseμ

Assuming that WSP is not defeated in word-final position, the resulting word-final heavy syllables of the winning candidates in (16b) and (17) will be stressed; there is no such requirement for the word-final light syllable of the winning candidate in (16a).

Elfner's analysis is ingenious, but there are, nevertheless, serious problems with it. First of all, it requires that word-final coronals be parsed as appendices connected directly to the prosodic word, when parsing the coronal as a coda would not create a superheavy syllable, e.g. *po.**e{t}*** of (16a). Elfner proposes that such an analysis follows from the ability of coronals to function as appendices to core syllable structure. However, arguing that a coronal consonant is an appendix when the coda position of the core syllable has not been filled is simply a contrivance to bring about the correct result with respect to syllable weight, and thus stress assignment to the word-final syllable – I am aware of no other evidence that this type of non-exhaustive syllabification is needed in English. While Elfner's analysis would nevertheless be reasonable if there was no other explanation of the observed vowel-reduction behaviour of the word-final syllables in (13) and (14), this not the case: the acoustic-perceptual analysis can provide a well-motivated, non-stipulative account. Notably, Elfner (2007) does not address Burzio's acoustic-perceptual analysis.

Second, and more seriously, Elfner's analysis sacrifices word-final consonant extrametricality. Hayes (1982: (24)) proposes that word-final consonants are extrametrical across all word classes in English. Later proposals (Giegerich, 1985; Burzio, 1994; Harris & Gussmann, 1998) similarly argue that word-final consonants are not syllabified as codas. Harris & Gussmann (1998) argue that such a generalisation can explain several phenomena of English stress and syllabification:

(i) Verbal stress contrasts: *édi<t>* versus *tormén<t>*.

(ii) Any consonant can appear after a long vowel word-finally in English, but not word-internally (Harris & Gussmann, 1998: (5)).

(iii) Closed syllable shortening, e.g. *fi<ve>~fif<th>*.

Elfner's analysis, like any stress analysis of the word-final syllables in question, requires that word-final consonants must sometimes be parsed as codas. The stress analysis therefore requires the sacrifice of a very useful generalisation about the metrification of word-final consonants in English.

There seems to be no such problem for Burzio's acoustic-perceptual account, because it is not reliant upon metrical structure. Myers (1987) points out that extrametricality refers to prosodic constituents, rather than the features associated with them:

> Crucially, it is only the extrametrical element itself that is inaccessible, and not the feature matrices associated with it. In *left*, for example, the final timing unit is extrametrical [...] but the features of the final /t/ still condition regressive voicing assimilation (cf. *leave*), since that rule is not sensitive to timing units (Myers, 1987: 487, f.n. 2).

Like the voicing assimilation Myers describes, the acoustic-perceptual analysis of vowel reduction is not crucially reliant upon syllable structure; it simply requires that vowels cue the consonants which are adjacent to them.[29] Final-consonant extrametricality does not prohibit this, as the features of an extrametrical final consonant will be still accessible and therefore able to condition vowel reduction – it is just the consonant's timing unit which is extrametrical.

In sum, there is no sound argument for syllables closed by velar or labial stops being more likely to receive stress than those closed by coronal stops. In an analysis which assumes a symmetrical relationship between vowel reduction and stress, it must be simply stipulated, as Ross (1972) is forced to do, that syllables with certain coda consonants are stressed, while others are not. Conversely, in an analysis such as Burzio's, where stress is not the only factor which can account for the

[29] Both in Burzio's own analysis and previously in Giegerich (1985), it is argued that word-final consonants constitute an onset: in Burzio's case, of a null vowel; in Giegerich's analysis, of a zero syllable. (This argument is also the standard in Government Phonology (Kaye, Lowenstamm & Vergnaud, 1985, 1990).) As both null vowels and zero syllables are acoustically empty, they cannot cue their consonant onsets, hence the job logically falls to the vowel preceding the word-final consonant (Burzio, 1994: 16).

absence of vowel reduction, the relationship between vowel reduction and obstruent quality is explained.

2.3.3 Summary and implications

It has been argued that the symmetrical relationship between vowel reduction and stress assumed in most standard generative analyses of English stress should be rejected. In §2.3.1, it was shown that an analysis which assumes a symmetrical relationship between vowel reduction and stress faces two undesirable complications to the stress system: the stressing of extrametrical syllables, and the acceptance of degenerate feet. In §2.3.2, it was shown that there is no need to assume a symmetrical relationship between vowel quality and stress and so accept the infelicities such an analysis incurs: Burzio's acoustic-perceptual account of vowel reduction offers explanations for why vowel reduction may be conditioned by coda quality directly, rather than relying upon stress as an intermediary. As well as avoiding the problems of stressing extrametrical syllables and introducing degenerate feet, the acoustic-perceptual analysis, unlike the stress analysis, can easily explain why final-syllable vowel reduction is conditioned by the place of articulation of word-final stops (§2.3.2.1).

The rejection of a symmetrical relationship between vowel reduction and stress has a huge implication for pre-tonic stress preservation – it removes the argument for it being stress that is preserved in pre-tonic position. As noted in §2.1 of this chapter, pre-tonic stress preservation has been argued for upon the basis of full vowel quality in the pre-tonic syllable alone. But, in light of the rejection of a symmetrical relationship between vowel reduction and stress, a full vowel in a pre-tonic syllable is no longer in itself confirmation of the presence of pre-tonic stress: this vowel may be full for independent reasons. With Burzio's theory in place, it would be an arbitrary and uneconomical stipulation to assume that pre-tonic syllables are stressed upon the basis of their full vowel quality alone.

It is important to note that the acceptance of Burzio's proposed asymmetrical relationship between stress and vowel reduction does not require the acceptance of the less desirable aspects of his theories. Notably, Burzio (1994) proposes that pre-tonic syllables cannot be stressed because this work absolutely rejects the monosyllabic heavy foot that would be created with pre-tonic stress: *(còn)**(dèn)**(sá)tion*.[30] Burzio's rejection of monosyllabic feet is problematic for two reasons.

The first problem with Burzio's total rejection of monosyllabic feet is that, as noted by Kager (1995b: 2), the monosyllabic (ˈH) foot has been argued for in many languages, and appears to pattern with (ˈLL) feet – rejection of the monosyllabic foot is therefore likely to be typologically implausible (see also Pater, 1995: 2, f.n. 2).

A second problem with Burzio's assumption of a minimally bisyllabic foot occurs specifically in English. There are words like *thìrtéen*, *Tòrbáy*, *ùpsét*$_A$ where it is necessary to assume that the initial syllable is stressed, but there is no adjacent unstressed syllable which can be incorporated to render this word-initial foot bisyllabic. Evidence for initial-syllable stress in these words comes from phrasal stress shift when they are in pre-modifying position: e.g. *thírtèen mén*, *Tórbày gúesthouse, úpsèt gírl*.[31] It is widely recognised that this type of postlexical rhythmic adjustment cannot shift stress onto a previously stressless syllable (for empirical support of this see Hammond (1999b: 352)). The only way Burzio can therefore get around this problem is ad hoc: postulating initial null vowels in order that these initial stresses satisfy his minimally bisyllabic foot requirement. This forces him in turn to amend his basic foot typology of (H σ) and (σ L σ) to also included (ϕ H), weakening the conciseness that makes his theory so attractive (Yamada, 1998: 233) (for further problems with Burzio's null vowel proposal see footnote 14, above, on catalexis).[32]

Fortunately, within OT, the acceptance of Burzio's rejection of pre-tonic stress preservation does not mean we need accept Burzio's proposal that all feet are minimally bisyllabic. The violable constraints used by OT mean that it is possible to model a grammar where the foot structure of *condensation* may indeed be *(còn.den)(sá)<tion>*, as required if this word's pre-tonic syllable is unstressed, but where the initial syllables of words like *Tòrbáy* consist of a monosyllabic (ˈH) foot – *(Tòr)(báy)*. Within OT, the constraint *CLASH-HEAD (Plag, 1999; Pater, 2000) – 'no

[30] See also Giegerich (1985) and Kager (1989) for theories which propose minimally bisyllabic feet.
[31] Examples from and stress-shift confirmed in Jones (2003) and Wells (2000).
[32] Giegerich (1985: 274) similarly notes that these stress shifts pose a problem for theories which absolutely exclude monosyllabic feet.

stress adjacent to the main stress' – could be ranked so that pre-tonic syllables are never stressed except in word-initial position, where these syllables would otherwise go unparsed. This is shown in (19):

(19) Proposed distribution of monosyllabic feet

Inputs: /*Torbay*/, /*condensation*/, /*Timbuctoo*/	PARSE-σ	*CLASH-HEAD	WSP
☞ a. (Tòr)(báy)		*	
b. Tor(báy)	*!		*
☞ a. (còn.den)(sá)<tion>			*
b. (còn)(dèn)(sá)<tion>		*!	
☞ a. (Tìmbuc)(tóo)			*
b. (Tìm)(bùc)(tóo)		*!	

As long as there is no constraint ranked above *CLASH-HEAD which requires pre-tonic stress word-medially – for example, WSP that requires every heavy syllable be stressed (*pace* Pater, 1995, 2000) – then it is possible to model the situation whereby pre-tonic monosyllabic feet are permitted in word-initial position, but not word-medially. In sum, I do not claim, like Burzio (1994), that monosyllabic feet should be excluded from the English foot inventory altogether. What is being proposed is that monosyllabic feet cannot be assumed to occur where their presence is argued for on the basis of the absence of vowel reduction behaviour **alone**; just such a situation is the argument for pre-tonic stress preservation in forms like *cònd**ĕn**sátion* (*condénse*).

2.4 Genuine pre-tonic stress preservation

The controversy over pre-tonic stress preservation presented in §2.3 resulted from the argument for pre-tonic stress preservation being motivated on the basis of vowel reduction behaviour alone. An issue not yet addressed is why the argument for pre-tonic stress preservation should have been so crucially reliant on observations about vowel reduction in the first place: i.e. why do we not get *expláin* → **explàinátion*, or *expéct* → **expèctátion*, where the second, preserving syllable is clearly the most prominent in the word after the primary stress? It is unlikely that such preservation would meet with the sort of controversy discussed in §2.3, as it would parallel the situation argued for left-edge preservation, e.g. *orìginálity* ← *oríginal*, where the presence of the preserved stress is not in doubt, thanks to its rhythmic salience.

In fact, genuine pre-tonic stress preservation does occasionally occur. Kager (1989: 171) argues that words which are candidates for pre-tonic preservation, and which have first and second syllables which are either both heavy or both light, occur consistently with initial-syllable secondary stress, e.g. *còndensátion*, *Jàpanése*. However, in words with light initial and heavy second syllables, pre-tonic stress preservation may occasionally occur, e.g. *acòustícian*, *depàrtméntal*, *elèctrícity*. These patterns are argued to be very good support for pre-tonic stress preservation, as the 0-2-1 pattern "is absent in segmentally parallel underived words" (Kager, 1989: 172).

The problem is that these examples of genuine pre-tonic stress preservation are highly uncommon, and are pretty much confined to the specific #LH phonological context. A search of Jones (2003) gives just the 14 examples in (20), out of a total of 221 words (all syllable weights) where pre-tonic preservation may be expected:[33]

(20) Apparently incontrovertible pre-tonic stress preservation

collèctívity (colléctive)	commèndátory (comménd)
connèctívity (connéctive)	detàinée (detáin)
detèstátion (detést)	dirèctórial (diréctor)
elàstícity (elástic)	elèctrícian (eléctric)
elèctrícity (eléctric)	ellìpsóidal (ellípse)
erùctátion (erúctate)	escàpée (escápe)
exchàngée (exchánge)	selèctívity (seléctive)

Most of these words also have initial-stressed patterns, e.g. *èlectrícity*, and potentially fall within Kager's #LH description based upon their dictionary pronunciations. I found no watertight examples of monomorphemic words with this

[33] I searched the Jones (2003) CD-Rom for words which have between 3 and 8 segments preceding the primary stress. Some of Kager's examples do not feature in (20), for example, *departmental*.

0-2-1 stress pattern, supporting the argument that the pre-tonic secondary stress in these words is due to preservation.[34]

Burzio (1994: 331; 2002: 2) argues that unmistakeable examples of pre-tonic stress preservation like **expèctátion*, **explàinátion* are generally unattested because such stress patterns generate a medial clash. Under the reasonable assumption that the prohibition against clash is responsible for the absence of **expèctátion*, we can infer the absence of pre-tonic stress in *èxp*[ɛ]*ctátion* or *cònd*[ɛ]*nsátion* – pre-tonic stress in the latter two examples would also generate a stress clash. We appear to have yet another piece of evidence against pre-tonic stress preservation in forms like *cònd*[ɛ]*nsátion*.

Burzio's argument is quite convincing, but it does rely upon one crucial assumption: that pre-tonic clash cannot occur to varying degrees. For example, the Rhythm Rule in Metrical Phonology (Liberman & Prince, 1977; Hayes, 1981) simply requires an alternating stress pattern - *thìrtĕen mén*, not *thĭrtèen mén* – rather than the complete absence of stresses on adjacent syllables (Liberman & Prince, 1977: 311-313). In the same way, one could argue that the absence of **expèctátion* does not automatically imply the absence of *èxpĕctátion*: if stress clash is gradient, then *èxpĕctátion* could be rhythmically more acceptable than **expèctátion*/*ĕxpèctátion*, in contravention of Burzio's argument. Nevertheless, as far as Pater's (1995, 2000) argument for pre-tonic stress preservation in OT is concerned, Burzio's argument stands: Pater (1995, 2000) cannot capture gradient word-internal rhythm for the pre-tonic context, as his analysis only employs *CLASH-HEAD, which prohibits stress clash at all levels.

Before concluding, we do need to consider how the few uncontroversial examples of pre-tonic stress preservation like *elèctrícity* may be formally handled. Interestingly, Kager (1989) suggests that there is some lexical restriction upon which #LH words can exhibit the pre-tonic stress preserving pattern:

> Although lexical variation may exist among words of the type [*annèxátion*] this variation should no be confused with optionality. First, words occur for

[34] I found four examples which could be argued to have bound roots, rather than words, as their immediate bases: *elèctrólysis*, *elèctrómeter*, *elèctrónic* and *relùctívity*. However, in all four cases the influence of words – *eléctric* and *relúctant* – seems intuitively likely.

> which no 3-4-1 contour exists, cf. *apartmental* [...] and second, words occur which show only a 3-4-1 contour, cf. *annexation* [...] Lexical variation and exceptionality go hand in hand for lexical rules (Kager, 1989: 174).

This lexical restriction of genuine pre-tonic stress preservation can be handled in more than one way in OT. Pater (2000: 259-260) proposes a lexically restricted preservation constraint that is ranked above PARSE-σ: in the words where this constraint applies, preservation occurs at the expense of parsing the word-initial light syllable, as in *a(pàrt)(mén)tal*, not *(àpart)(méntal)*. In line with the Stratal OT analysis which I propose in chapter 8 of this thesis, an analysis in terms of cophonologies (Anttila, 2002a) (see §9.3.2), rather than lexically indexed constraints, would be more appropriate.

2.5 Probabilistic pre-tonic preservation: Hammond (2003a, b)

I have dismissed the argument for pre-tonic **stress** preservation in words like *cond*[ɛ]*nsation*, but I do accept that there is pre-tonic vowel quality preservation. In light of this proposal for preservation, it is worth presenting the frequency analysis given in Hammond (2003a, b) which attempts to account for the inconsistency of pre-tonic preservation within the #HH(son coda) context, e.g. *cònd*[ɛ]*nsátion* (*cond*[ɛ́]*nse*) versus *ìnf*[ə]*rmátion* (*inf*[ɔ́ː]*rm*) (see §2.1 and §2.2).

Hammond (2003a, b) proposes that pre-tonic vowel quality in embedding words is influenced by both the frequency of the embedded word from which phonology is preserved, and the frequency of the embedding word itself. [35] Hammond's observation that phonological identity effects may be affected by word frequency is important because, in chapter 6, it will be shown that left-edge stress preservation (e.g. ***àntí**cipate* → ***ăntì**cipátion*) is also an example of a probabilistic phonological identity effect. Here, Hammond's frequency analyses are briefly reviewed, as they provide a starting point for my own frequency analyses. Some

[35] Hammond takes the traditional approach and argues that a non-reduced vowel is an indicator of stress, but his results can equally be understood as preservation of vowel quality directly, as it is the behaviour of vowel reduction which he examines.

potential problems with Hammond's conclusions are also noted; these are addressed in more detail in chapter 6.

2.5.1 Hammond's frequency analysis of pre-tonic vowel reduction

Fidelholtz (1975: 200) shows that vowel reduction in word-initial closed syllables is affected by frequency: the relatively frequent word [ə]*strónomy* undergoes initial reduction more readily then the relatively infrequent word *g*[æ]*strónomy*. This is in line with the generally accepted observation whereby high frequency forms tend to be perceptually more ambiguous than low frequency forms (Hay, 2003: 5).

Hammond (2003a) tests to see if Fidelholtz's generalisation about frequency and vowel reduction extends to word-internal pre-tonic syllables. Hammond explores not only the effect of the frequency of the embedding word itself upon pre-tonic vowel reduction, but crucially, with respect to preservation, whether the frequency of the embedded word has an effect of increasing or decreasing the likelihood of reduction in the pre-tonic syllable of the embedding word. Therefore, for a word like *condensation*, the frequency of *condense* as well as that of *condensation* is examined.

Hammond's data is a list of -*ation* nouns and their embedded words collected from the MRC Psycholinguistic database, along with the pre-tonic vowel qualities of the embedding word.[36] The chosen sample contains embedding words with second syllables closed by sonorants only, in light of effects of coda quality upon vowel reduction discussed so far in this chapter. Hammond's token frequencies for embedding and embedded words are based upon the number of occurrences in the Brown Corpus. Pater (2000) argues against preservation for words with three pre-tonic syllables, e.g. *recommendation* (see f.n. 7), but these examples are not excluded from Hammond's analysis.

In regression analyses between embedded and embedding frequency and vowel quality (pre-tonic reduced or unreduced), Hammond shows significant effects upon vowel reduction behaviour for both embedding and embedded word

[36] Hammond (2003a) does not address instances of variable pre-tonic vowel quality, although Pater (2000) argues that there is speaker-internal variation. From my own examination, the MRC Psycholinguistic database appears to give consistent pronunciations one way or the other only, in contrast to the more often variable pronunciations given by Jones (2003) and Wells (2000). Hammond (2003b, 2004) takes data from *Websters' New Collegiate Dictionary*, and deals with variable pre-tonic vowel quality also. The results in this analysis are significant, like those given here.

frequencies: $R^2 = 0.3441$, p= 0.0002 for the frequency of the embedding word; $R^2 = 0.1363$, p= 0.0267 for the frequency of the embedded word.

As is anticipated in light of Fidelholtz's original study, Hammond reports that vowel reduction in pre-tonic syllables of embedding words is more probable when the embedding word is more frequent. What is more surprising is Hammond's claim that the probability of vowel reduction increases as the frequency of the embedded word increases. Hammond (2003b: 10) himself notes that it is not what we might anticipate: we would intuitively expect vowel reduction to become less likely the more frequent, and therefore the more salient, the embedded word is. Certainly, in light of the frequency effects proposed by Hay (2001, 2003), Hammond's conclusion is unexpected: as will be discussed at length in chapter 6, Hay proposes an effect of relative frequency, such that we might expect vowel reduction to become more likely the less frequent the embedded word is relative to the embedding word. In contrast, Hammond proposes that we are witnessing an effect of cumulative frequency, which he claims is supported by research such as Taft (1979) (Hammond, 2003a: 12): the greater the combined frequency of the embedding and embedded words, the more likely vowel reduction is in the embedding word.[37]

Hammond's argument for an effect of cumulative rather than relative frequency seems particularly counterintuitive in light of a pilot study by Kraska-Szlenk (2007: 142-3). Kraska-Szlenk examines words where pre-tonic preservation is expected, along with their own frequencies and those of their embedding words. The 15 embedding words Kraska-Szlenk examines are from Pater (2000); the frequency information is taken from the British National Corpus. The data is given in (21); numbers in brackets indicate the number of times a word occurs in the British National Corpus:

[37] Admittedly, it is not impossible to conceive of a situation whereby both cumulative and relative frequency effects could be observed with the same set of data: the data set could consist of high frequency embedding words, all of which happen to be more frequent than their also high frequency bases, as well as low frequency embedding words, all of which happen to be less frequent than their also low frequency bases. However, as I show in chapter 6, no conclusions can be made on this basis with Hammond's data, and there is no observable cumulative frequency effect with my stress preservation data.

(21) Data from Kraska-Szlenk (2007)

(a) Reduced pre-tonic vowel

(i)	ìnf**or**mátion (38327)	inf**ór**m (286)	
	cònv**er**sátion (5169)	conv**ér**se (13)	
	còns**er**vátion (3943)	cons**ér**ve (55)	
	còns**ul**tátion (2593)	cons**úl**t (319)	
	trànsf**or**mátion (1712)	trànsf**ór**m (187)	
	cònf**ir**mátion (1144)	conf**ír**m (751)	
	trànsp**or**tátion (553)	trànsp**ór**t (21)	
(ii)	sègm**en**tátion (247)	sègm**én**t (0)	(cf. ségm**en**t (749))
	pìgm**en**tátion (54)	pìgm**én**t (0)	(cf. pígm**en**t (223))

(b) Full pre-tonic vowel

(i)	àdv**àn**tágeous (372)	adv**án**tage (7220)
	àuth**èn**ticíty (362)	àuth**én**tic (824)
(ii)	cònd**èm**nátion (443)	cond**ém**n (473)
	cònd**èn**sátion (336)	cond**én**se (61)
	ìmp**òr**tátion (164)	imp**ór**t (62)
	àugm**èn**tátion (53)	àugm**én**t (19)

For the non-preserving examples in (21a, i), the frequency of the embedding word is massively greater than the frequency of the embedded word. Non-preservation also occurs without massively high embedding frequency in (21a, ii), but in these cases, there may be interference from the bracketed high frequency nouns *ségment* and *pígment*. With preservation in (21b), the embedding word is not so greatly frequent compared to the embedded word, cp (21a, i). For the examples in (21b, i), the embedded word is much more frequent than the embedding word. For the examples in (21b, ii), the embedded word is less frequent than the embedding word, but the difference is much smaller than the massive differences observed in the examples in (21a, i). Therefore, while Kraska-Szlenk's results are impressionistic, they seem to be in support of Hay's hypothesis for relative frequency, rather than Hammond's for cumulative frequency. In particular, the case of *àdv**àn**tágeous* (372) - *adv**án**tage*

(7220) is particularly problematic for Hammond's cumulative frequency proposal: we might expect the massive frequency of the embedded word to induce vowel reduction in the embedding word, but this does not occur.

Further time will not be spent debating the correctness of Hammond's frequency analysis and conclusions here, as these issues are addressed further in §6.1.1. For now, this discussion of Hammond's analysis of pre-tonic preservation raises an important issue that that will resurface in chapters 6 and 8 of this thesis: the effect of word frequency upon morphophonological identity between words.

2.6 Conclusion

In this chapter, it has been argued that there is no evidence for pre-tonic stress preservation in English in words with initial-syllable secondary stress, e.g. *còndensátion*. The case for this pre-tonic stress preservation relies crucially upon the assumption that a symmetrical relationship exists between vowel reduction and stress; in §2.3, I showed that such an assumption is both problematic and unenlightening when compared to Burzio's (1994, 2002) acoustic-perceptual analysis of vowel reduction which rejects a symmetrical relationship between vowel quality and stress. By rejecting the assumption that there is a symmetrical relationship between pre-tonic vowel quality and stress, we have no evidence that it is a low level of stress, rather than just vowel quality directly, that is preserved in pre-tonic position.

While I have made a strong case for rejecting pre-tonic stress preservation in this chapter, there is a notable omission of an analysis of pre-tonic vowel quality preservation to take its place. At the time of writing, no such analysis has been proposed in any depth by Burzio. Such an analysis will not be undertaken in any detail in this thesis: my primary aim is to propose an in-depth formal handling of English weak stress preservation, rather than more diverse morphophonological identity effects between words. In the conclusion of the thesis, it will be noted that probabilistic pre-tonic vowel quality preservation is an area for future research.

We must turn elsewhere, therefore, for indisputable examples of weak stress preservation in English. In chapter three, I review the case for left-edge stress preservation – e.g. *oríginal* → *orìginálity*, *àntícipate* → *ăntìcipátion* – in English, and show that this second type of weak stress preservation is also not without some controversy.

Chapter three: left-edge stress preservation

3.0 Introduction

In this chapter, the status of a second type of weak stress preservation that has been argued to occur in English is examined: left-edge stress preservation. The argument for left-edge stress preservation does not face the type of controversy observed for putative pre-tonic stress preservation in the last chapter, where it was not clear that it was stress which was being preserved. Nevertheless, it is concluded that the status of left-edge stress preservation is still somewhat controversial, and that further empirical investigation is required to establish its existence and precise nature.

As outlined in the introduction to the thesis, left-edge stress preservation is a term used here to denote weak stress preservation that occurs upon one of the first two syllables of an embedding word, where the embedding word has more than two pre-tonic syllables. Fully successful left-edge stress preservation results in the preserved stress being the second-most prominent in the word: e.g. *oríginal* → *orìginálity* (0-2-0-1-0-0). The syllable with preserved stress is rhythmically more prominent than either of its unstressed neighbours, as shown in (1):

(1) Left-edge stress preservation: rhythmic salience

```
         x
   x     x
x  x  x  x  x  x
o  ri gi na li ty
```

This is quite a different situation to putative pre-tonic stress preservation, (2):

(2) Putative pre-tonic preservation: rhythmic weakness

```
          x
x         x
x    x    x
x    x    x  x
con  den  sa tion
```

Any putatively preserved stress in *cònd**ĕn**sátion* is the weak member of a rhythmic alternation, being flanked on either side by rhythmically more prominent stressed syllables. It is this lack of rhythmic salience which had led to the situation whereby vowel reduction behaviour has been invoked to determine the presence or absence of pre-tonic stress (chapter 2).

It can be argued that there are two sub-types of left-edge stress preservation. The first to be addressed in this chapter, and the most controversial, is what will be referred to as 'relative prominence preservation'. The argument for relative prominence preservation comes from Kiparsky (1979), who claimed that the relative prominences of feet are preserved under morphological embedding: *sĕns**à**tionálity* ← *sèns**á**tional*. The second sub-type of left-edge stress preservation to be addressed in this chapter is foot-head preservation. With foot-head preservation, only one of the first two syllables of an embedded word is stressed, and so simply preserving the locations of foot-heads will ensure the correct stress contour in the embedding word: *or**ì**ginálity* ← *or**í**ginal*. We will see in this chapter that foot-head preservation is also not totally uncontroversial.

This chapter is structured as follows. In §3.1, the placement of left-edge stress in words which are not candidates for preservation is discussed. In §3.2, I examine the previous arguments made for and against relative prominence preservation, before presenting the somewhat less controversial case of foot-head preservation in §3.3. The differences in stress behaviour that are likely to be observed between words which are candidates for left-edge preservation, and those which are not, are summarised in §3.4; any differences would provide support for left-edge stress preservation, either of foot-heads or of relative prominence, in the empirical investigations which follow this chapter. Finally, in §3.5, the chapter concludes with a summary of the areas that require further research in order to establish the precise status of left-edge stress preservation.

Before proceeding, a disclaimer is required: to keep my empirical investigation in this thesis to a manageable size, I consider left-edge preservation only in the context of three pre-tonic syllables. Previous and much less exhaustive examinations of the left-edge stress preservation data made by Halle & Vergnaud (1987a) and Sainz (1988, 1992) have included words with more than three syllables preceding the primary stress, e.g. *irrèparabílity* (4 syllables). In a careful investigation, each phonological context will need to be considered separately, as phonological and morphological factors interact to predict if and how stress preservation manifests. I leave the examination of longer sequences of pre-tonic syllables to future work.

3.1 Left-edge stress assignment without preservation

In this section, the assignment of left-edge stress to words which are not candidates for stress preservation is discussed. This is important because, if words which are candidates for left-edge stress preservation differ from these default patterns in terms of their stress placement, it will provide evidence for left-edge preservation.

It is evident from much previous literature that the placement of non-primary stress in English words is not always determined by alternating rhythm alone. In the environment of three pre-tonic syllables being discussed here, it is recognised that the assignment of non-primary stress is also weight-sensitive.[1] In some cases, this weight-sensitivity can determine whether the first or second syllable of a word is stressed, as in ***à**bracadábra* versus *Mon**òn**gahéla* (SPE: 114; Halle & Keyser, 1971: 50; Liberman & Prince, 1977: 276; Hayes, 1982: 258-259; Sainz, 1988, 1992; Kager, 1989: 43-44; Pater, 1995, 2000; Hammond, 1999a: 295). Weight-sensitivity has also been argued to determine the stressing of word-initial syllables: where the first syllable of a word as well as the second syllable is heavy, e.g. *T*[aɪ/iː]*conderoga*, its initial syllable will be stressed, in addition to its second syllable – ***Tì**cònderóga* (SPE: 118, f.n. 72; Halle, 1973: 460; Kiparsky, 1979; Selkirk, 1980: 571; Hayes, 1981: 170, 1982: 260; Sainz, 1988, 1992; Hammond, 1999a: 295-6), cp. light-initial ***Mo**nòngahéla*.

In (3), the stress-assignment behaviour that is expected in the context of three pre-tonic syllables, depending upon the particular combination of syllable weights involved, is consolidated. This summary is based particularly upon Kager's (1989:

[1] Of course, it is not the case that English secondary stress is totally weight-sensitive – rhythmic principles may sometimes override weight-sensitivity, e.g. *Àlexánder* (heavy second syllable unstressed). Pater (1995, 2000) gives a particularly good account of this weight 'non-uniformity'.

44) useful overview, but reflects the observations of the diverse literature cited in the previous paragraph (the examples are also taken from this range of literature).

(3) Secondary stress assignment with three pre-tonic syllables

H = heavy syllable; L = light syllable; # = word boundary

(i) The initial syllable is stressed if the second syllable is light, regardless of the weight of the immediately pre-tonic syllable:

#LLL: ***à****bracadábra* #LLH: ***Kì****limanjáro* #HLL: ***Lùx****ipalílla*

Kà*lamazóo* ***Hà****licarnássus* ***Hàr****decanúte*

(ii) Heavy second syllables are always stressed:

#HHL: *Tì****còn****deróga* #LHL: *Mon****òn****gahéla* ([mə.nɒŋ]*gahela*)

*Òmp****òm****panóosuc* *Bel****ù****chistán* ([bɛ/ə.luː]*chistan*)

(iii) When the second syllable is stressed, initial heavy syllables are stressed, but light ones are not:

#HHL: ***Tì****cònderóga* #LHL ***Mo****nòngahéla*

It will have been noted that (3) is not exhaustive. With three pre-tonic syllables, there are 8 (2^3) different possible sequences of syllable weight, but only 5 are discussed in (3). In literature which addresses trisyllabic pretonic sequences, there is very little evidence indeed in support of the three remaining sequences: #HHH, #LHH and #HLH. The only uncontroversial example of #LHH given in the literature cited above is *Tenòchtitlán* [tə.nɒk.tɪt-] (Hammond, 1999a: 295), whose stress pattern does fit in with the generalisations in (3). No uncontroversial examples of #HHH have been found in the literature.[2] It will be seen during the rest of this chapter and the following two that these two sequences of syllable weight are also very rare or absent in words that are candidates for left-edge stress preservation. There is also an absence of discussion of #HLH examples which are not candidates

---

[2]*Àlexandrétta* (also *-inus*) is given by Hammond (1999a: 297), but it is possible that the stress of *Álex* or *Àlexándra* interferes with this. I have found just one apparently sound example of either pattern in my dictionary research (detailed in the following chapters): #LHH and #HHH (variable) in *Bophuthatswana*, from Jones (2003). According to Jones, this word has the stress pattern *Bòphuthatswána*.

for preservation in the literature I have examined,[3] although it will be noted in the next chapter that this pattern is reasonably frequent in complex words where initial-syllable preservation is anticipated.

An area of ambiguity is the stressing of the second syllable of a pre-tonic #HLL sequence – words like *Luxipalilla* and *Dodecanesian*. The conditions in (3) do not predict whether or not this second syllable is stressed. Kiparsky (1979: 423) argues that both the first and second syllables of #HLL *Dodecanesian* are heads of feet.[4] However, Hammond (1989: 145), Halle & Kenstowicz (1991: 490) and Hayes (1981: 149) all indicate that light second syllables are not generally the head of a foot: both Hammond and Halle & Kenstowicz argue for contrasts between monomorphemic *Lùxipalílla* and candidates for second-syllable preservation like *ànt**ì**cipátion* and *ìc**ò**noclástic*. Hammond argues that monomorphemic examples like *Dòd**è**canésian* do have stress on their second syllable, but that this is exceptional.[5] In chapter 4, it will be seen that Hammond's proposal appears to be supported, although in chapter 9 we will see that this causes some problems for Pater's (1995, 2000) handling of English secondary stress.[6]

In table 1, the left-edge stress behaviour we expect in words with three pre-tonic syllables is summarised.

---

[3] Again, examples are given just by Hammond (1999a: 296), who has carried out an extensive search. However, Hammond's examples do not transfer into British English pronunciations (e.g. long General American [ɔː] versus short British English [ɒ]) and/or his syllable-weight classifications are questionable (e.g. the initial syllable of *Buenaventura* – [bwɛ] – as heavy).

[4] There is some question over whether Kiparsky's example *Dodecanesian* is in fact monomorphemic, as he implies: Jones (2003) gives *Dòdecanése*. This does not cause problems for the argument here, as there is no stress on the second syllable of *Dòdecanése* – *Dodecanesian* is still not a candidate for second-syllable stress preservation.

[5] The light second syllable is marked as an exception to Hammond's destressing rule that destresses subsidiary-stressed light syllables adjacent to a heavy stressed syllable (Hammond, 1989: 146-149).

[6] One could speculate a couple of reasons why the pre-tonic footing #(H)(LL) could be preferred to #(HL)L in a language: (i) the former parsing avoids unparsed syllables; (ii) the latter parsing creates a non-canonical moraic trochee (HL) (on canonical and non-canonical feet see Kager (1993), Hayes (1995) and Alber (1997)).

Syllable weight			Stress pattern
Initial syllable	Second syllable	Third syllable	
L	L	L	**à**bracadábra L̀LL
L	L	H	H**à**licarnássus L̀LH
L	H	L	Mon**òn**gahéla LH̀L
L	H	H	___
H	L	L	L**ù**xipalílla H̀LL (D**ò**d**è**canésian) H̀L̀L
H	L	H	___
H	H	L	T**ì**c**òn**deróga H̀H̀L
H	H	H	___

Table 1: secondary stress expected in monomorphemic words

In the rest of this chapter, we will see how the stress behaviour predicted by left-edge stress preservation differs from this control behaviour.

3.2 Relative prominence preservation

The original argument for relative prominence preservation comes from Kiparsky (1979). Kiparsky argued that, in words with three pre-tonic syllables where both the first and second syllables are stressed, there should be variation in whether the stress of the initial or second syllable is more prominent when the word is not a candidate for relative prominence preservation (see also Selkirk, 1980: 601-2; Kager, 1989: 64, 168-70; Sainz, 1992: 90):

(4) Variable prominence in heavy-initial monomorphemic words

Tĭc**òn**deróga~T**ì**cŏnderóga *Tĭc**ón**dero

Sr̆ir**àn**gapátnam~Sr**ì**răngapátnam *Srir**án**gapat

Kiparsky proposed that there should be no such choice in parallel complex words where preservation occurs: i.e. only *sĕns**à**tionálity* should arise from *sèns**á**tional+ity*, never ***sèn*** *sătionálity* (Kiparsky, 1979: 423).[7]

Kiparsky's argument for relative prominence preservation is a response to Liberman & Prince (1977). Following work like Schane (1975), Liberman & Prince distinguished between the simple presence or absence of stress, and the relative prominence of these stresses. Like Schane, Liberman & Prince reject the cycle as the means by which stress is subordinated, instead proposing that the very presence of stress is indicated by a segmental diacritic [+stress], but the relative prominences of different degrees of stress are captured using using metrical trees. Liberman & Prince (1977: 301) proposed that metrical trees below the word level were not cyclically preserved, but were rather erected anew on each cycle – 'Deforestation'. Only the markers for the locations of foot-heads – [+stress] – were cyclically preserved.

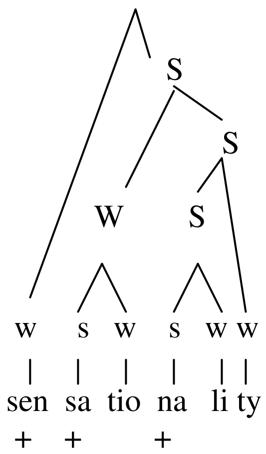
Liberman & Prince's proposal for Deforestation meant that their theory could not predict within-word relative prominence preservation in *sĕnsàtionálity*. Under Kiparsky's argument for relative prominence preservation, metrical structure, as well as [+stress], had to be cyclically preserved: if only [+stress] was carried over, then "one could equivalently stipulate that metrical structure is assigned only on the last cycle" (Kiparsky, 1979: 422). The problem with assigning metrical structure on the last cycle in this way was that it predicted an indeterminate relative prominence contour over the first two syllables of a word like *sensationality*, shown in (5).

[7] Kiparsky's argument absolutely relies upon the differentiation between the prominence of non-primary stresses. In more recent generative accounts of word stress, e.g. Halle & Vergnaud (1987a), Burzio (1994), Pater (1995, 2000), Hammond (1999a), the tendency is to simply treat all subsidiary stresses as equivalent in prominence. Burzio (1994: 186) takes such an approach, and can only account for the rising prominence contour over the first two feet of what is *sènsàtionálity* in his analysis by using his rather problematic word-initial weak syllables (re: chapter 2, f.n. 13). Therefore, like Kiparsky, I differentiate between levels of subsidiary prominence.

(5) Variable relative prominence in Liberman & Prince (1977)

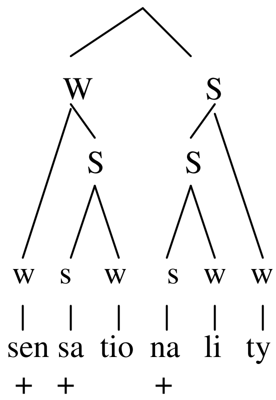
\+ Liberman & Prince (1977)'s [+ stress]

(a) Right-branching: 2-3-1



Liberman & Prince (1977)

(b) Left-branching: 3-2-1



Liberman & Prince (1977) and Kiparsky (1979)

In contrast, under Kiparsky's proposal for the preservation of metrical structure, the metrical tree of the embedded word *sènsátional* would be preserved, ensuring only a left-branching tree and so the correct 3-2-1 stress contour in the embedding word.

Since Kiparsky (1979), arguments have been made against the existence of relative prominence preservation which vindicate Liberman & Prince's argument. These are now discussed in §3.2.1.

3.2.1 Arguments against relative prominence preservation

Halle & Vergnaud (1987a: 242-6), Kager (1989: 170) and Sainz (1992: 90) all argue against relative prominence preservation on the basis of its failure in a number of complex words. All works cite examples of complex words that do not consistently preserve the relative prominence contour of their base, based upon pronunciations given in Kenyon & Knott (1944/53); some examples are given in (6).

(6) Failure of relative prominence preservation

ăccèlerátion~àccĕlerátion	(#HL)	(àccélerate)
ăccèntuátion~àccĕntuátion	(#HH)	(àccéntuate)
ăntìcipátion~àntĭcipátion	(#HL)	(àntícipate)
hŭmànitárian~hùmănitárian	(#HL)	(hùmánity)
hŭmìliátion~hùmĭliátion	(#HL)	(hùmíliate)
ĭcònoclástic~ìcŏnoclástic	(#HL)	(ìcónoclast)
ĭmmùtabílity~ìmmŭtabílity	(#HH)	(ìmmútable)
ĭmpèccabílity~ìmpĕccabílity	(#HL)	(ìmpéccable)
ĭnfàllibílity~ìnfăllibílity	(#HL)	(ìnfállible)
ĭnfèriórity~ìnfĕriórity	(#HH)	(ìnférior)
tŏtàlitárian~tòtălitárian	(#HL)	(tòtálity)
trănsfigurátion~trànsfĭgurátion	(#HL)	(trànsfígure)

Rather than consistently preserving the relative prominence contour of their embedded word by having secondary prominence upon their second syllables, the embedding words in (6) sometimes have secondary prominence upon their initial syllables. They are therefore behaving the way that Kiparsky (1979) argued only words which are not candidates for preservation, like *Ticonderoga*, should.

In both Halle & Vergnaud (1987a) and Sainz (1992), the variable relative prominence contours given in (6) are captured by variable rule application. Halle & Vergnaud capture this variable relative prominence with their rule of Stress Enhancement, (7), which assigns a line 2 asterisk to either of the first two feet of embedding words like those in (6):

(7) Halle & Vergnaud's Stress Enhancement (1987a: 242)

*	line 2
* → * / [(SYL) ___	line 1

Sainz (1992: 94) captures this variable relative prominence through her Rhythm Rule, (8), and her Peninitial Stress rule, (9):

(8) Rhythm Rule (Sainz, 1992: 94)

```
          x                 x        row 3
      x   x         x       x        row 2
  x   x   x   →     x   x   x        row 1   e.g. Tìcŏnderóga
```

(9) Peninitial Stress (Sainz, 1992: 91)

```
                      x               row 2
   x   x            x x               row 1
  [(x)(x   x)  →  [ (x)(x   x)        row 0   e.g. Tĭcònderóga
```

The Rhythm Rule assigns greater prominence to the initial syllable of a trisyllabic pre-tonic sequence where both initial and second syllables are heads of feet.[8] Sainz's Rhythm Rule is motivated by the principle of eurhythmy that is captured in Hayes's (1984: 46) Quadrisyllabic Rule – a grid should have a row whose grid marks are ideally four syllables apart. Sainz's Peninitial Stress rule accounts for variants where the second syllable is more prominent than the first, and is akin to Halle & Vergnaud's rule of Stress Enhancement in its ad hoc nature. Sainz's Rhythm Rule effectively corrects the outputs of her Peninitial Stress Rule, as shown in (10):[9]

(10) Action of Sainz's Rhythm Rule

```
(a) Output of Peninitial Stress          (b) Output of Rhythm Rule

        x                                        x        row 3
    x --- x                               x ----- x        row 2
  x x     x          Rhythm               x x     x        row 1
  x x  x x x          Rule                x x  x x x       row 0
  Ticonderoga          →                  Ticonderoga
```

[8] Sainz (1992) does not discuss how the light second syllables of words like *iconoclastic* and *Dodecanesian* may get stressed.

[9] By capturing initial-prominent patterns using the Rhythm Rule, and with no independent motivation for the Peninitial Stress rule, Sainz's analysis implies that the initial-prominent variants are more well-formed than second-syllable prominent patterns. Selkirk (1980: 600-1) argues that this is indeed the case. In contrast, neither Kiparsky's (1979) argument nor that of Halle & Vergnaud (1987a) indicates that a 3-2-0-1 pattern is any less well-formed than a 2-3-0-1 pattern. In what is effectively a revision of Halle & Vergnaud's (1987a) Stress Enhancement, Halle & Kenstowicz (1991: 492) argue that initial-syllable prominence is the "unmarked" case.

It is important to note that the analyses which predict variable relative prominence do not claim that every eligible word should occur with both stress patterns. Both Sainz and Halle & Vergnaud propose that the variation may be lexically restricted, and Kiparsky (1979: 423-4) makes the same argument with respect to words which are not candidates for preservation. In Sainz's analysis, words are marked as to whether the Rhythm Rule is prohibited or applies optionally or obligatorily. Halle & Vergnaud similarly propose that Stress Enhancement is lexically restricted, and argue that if a word does not have both variants, this is an "oversight" rather "than a systematic gap to be accounted for in a phonological description of a language" (Halle & Vergnaud, 1987a: 245). It is therefore not automatically damning for the analyses which reject relative prominence preservation that that there are words where only the prominence preserving pattern occurs, e.g. *pŏstèriórity*, *sŏmnàmbulátion*, *sŭpèriórity* (Kager, 1989: 170). It may be argued that these words just happen to have only this variant.

Nevertheless, if the absence of variation in forms like *pŏstèriórity* does not automatically negate the argument against relative prominence preservation, nor does the variation in complex words like *ìcŏnoclástic* ~ *ĭcònoclástic* automatically negate the argument for relative prominence preservation. It is perfectly plausible that the true state-of-affairs lies somewhere in the middle: relative prominence preservation occurs, but it is not one-hundred percent successful. One possible explanation for such a situation is that morphological complexity is gradient rather than absolute. Such was the argument made for pre-tonic vowel quality preservation in §2.5, and is the line of inquiry pursued in the following chapters.

There is also some doubt over the quality of the data presented by Sainz, Halle & Vergnaud, and Kager. The counterexamples to relative prominence preservation are taken from Kenyon & Knott (1944/53), where the possible explanations for the documented variable relative prominence in the embedding words are as follows:

> In actual speech, such alternative accentuations such as *a,cade'mician* or *,acade'mician*, *,impenetra'bility* or *im,penetra'bility*, *in,feri'ority* or *,inferi'ority* are very common, and do not represent more or less desirable pronunciations, but chiefly show the effect of varying sense stress, emphasis, speech rhythm, semantic distinctions, and other constantly varying factors of

> connected speech, so that in many such instances the question which accentuation is preferable is irrelevant (Kenyon & Knott, 1944/1953: xxv [Kager, 1989: 170; see also Halle & Vergnaud, 1987a: 243-4]).

(Kenyon & Knott's example *academician* does not fit in with the description of words which vary as having a heavy initial syllable, and will be returned to shortly.) From this quotation, it is clear that one must be cautious when using Kenyon & Knott's data to suggest that there is never relative prominence preservation. Stress preservation is argued to apply at the Stem and Word levels, rather than the Phrase level, and some of the factors given by Kenyon & Knott as causing variation are 'factors of connected speech': variable prominence may simply be a phrasal phenomenon that overrides consistent relative prominence preservation at a lower level. This possibility will also be entertained in later chapters.

A third query for the arguments which completely reject relative prominence preservation is Kager's (1995b) suggestion that variable relative prominence in the embedding word may be the **result** of preservation, rather than the failure of it:

> It is interesting to observe that the 'immediate sub-constituent principle' does not always predict transfer of stress variability to the derived word. For example, Kenyon & Knott (1953) note stress variation in, e.g., *imprègnabílity*, *ìmpregnabílity*, arguably corresponding to the stress variation *imprégnàte* and *ímpregnàte* (the former is favored in American English, the latter in British English). But notice that there is no variation in the immediate base of the word, which is *imprégnable* (**ímpregnable*) (Kager, 1995b: 21).

Admittedly, Kager's proposal here requires some refinement given its suggestion that the pronunciation of one variety of English may be influencing the pronunciation of the other. Nevertheless, it does highlight the fact that we need to be very careful about what we class as stress preservation failure, particularly as the same point is made by Marvin (2002: 71) with respect to relative prominence preservation: Marvin attributes variation in *ăccèptabílity~àccĕptabílity*, recorded by Kenyon & Knott, to two possible pronunciations of the more-deeply embedded word *accept*. The nature of the phonological cycle has led research to focus upon the inheritance of stress from immediately-embedded constituents, but, in the face of the theories of Output-Output Correspondence that have more recently been proposed in Optimality Theory (chapter 7 of this thesis), as well as the theory of fake cyclicity proposed in chapter 8, wider-ranging phonological relationships should be entertained. Therefore, any argument which wishes to use examples of variable relative prominence as examples of the failure of relative prominence preservation must verify that the variation in the embedding word could not be attributed to preservation from another source. This has not yet been attempted systematically in any literature, and so will also form a part of the empirical investigation in the following chapters.

In sum, although data has been presented which certainly questions the existence of relative prominence preservation, this evidence is by no means conclusive. None of the works reviewed in this section carry out the sort of detailed quantitative research which is required to establish the non-existence or otherwise of relative prominence preservation. Both Halle & Vergnaud (1987a: 245-6, f.n. 12) and Sainz (1992: 129-30) list just over one hundred words from Kenyon & Knott (1944/53) with variable secondary stress, but this list includes a whole range of syllable weight combinations, and is not compared to the number of complex words that occur with the consistently preserving pattern. New quantitative research is required to establish which of situations (i)-(iii) is, in fact, correct:

(i) Variation is due to across-the-board relative prominence preservation failure, as proposed by Halle & Vergnaud (1987a), Kager (1989) and Sainz (1992).

(ii) A later rhythmic adjustment is overriding the outputs of successful relative prominence preservation, the possibility of which is implied by Kenyon & Knott (1944/53).

(iii) There is relative prominence preservation in some instances, but the phenomenon is variable in its success – a new hypothesis proposed here.

In chapter 4, I carry out the necessary research to establish which of outcomes (i)-(iii) is the case, and show that the results point towards (iii).[10] This result should not really be surprising: we already saw in chapter 2 that pre-tonic vowel quality preservation may well be probabilistically dependent upon word-frequency effects. However, for this argument to make sense, it must be shown that variable secondary-

[10] However, as shown in §6.5.4.2, (ii) and (iii) are not mutually incompatible.

stress placement is not limited to words which are candidates for relative prominence preservation, but also occurs in words which are candidates for just left-edge foot-head preservation (*oríginal* → *orìginálity*). If stress preservation failure can fail as a function of morphological decomposability, then it should affect all types of preservation. In the following chapters, we will see that this seems to be the case.

3.3 Left-edge foot-head preservation

Kiparsky's argument for relative prominence preservation is only relevant to words with heavy initial syllables.[11] This is because relative prominence preservation is only applicable to words where both the first and second syllables are footed, and the initial syllable of a word cannot be footed in addition to the second syllable if it would create a degenerate foot: **(ò)(rígi)nal*. When the first and second syllables of a word are light, only the first **or** second syllable is argued to be footed, depending upon whether or not the word is a candidate for stress preservation: *o(**rì**.gi)(ná.li)<ty>* ← *o(**rí**.gi)<nal>*, versus *(**à**.bra)ca(dá.bra)*. Relative prominence preservation is redundant here: simply carrying over the locations of feet to the next cycle will ensure preservation, as only the binary contrast of stressed versus unstressed is needed to predict preserving *o**rì**ginálity* (*o**rí**ginal*) rather than non-preserving *****ò**riginálity*.

There is greater consensus in the literature over the existence of foot-head preservation than over that of relative prominence preservation. Liberman & Prince's (1977) analysis predicts foot-head preservation even though it rejects relative prominence preservation: foot-head preservation requires only that the locations of foot heads are preserved from one cycle to the next, and this is ensured by the preservation of [+stress] which Liberman & Prince's analysis permits (§3.2, above). Similarly, although Halle & Vergnaud (1987a) reject relative prominence preservation, their rule of Stress Copy ensures that foot-head preservation occurs (see also Halle & Kenstowicz, 1991: 490-491). In more recent literature, the existence

---

[11] The stress rules in Halle & Vergnaud (1987a) make light-initial words like *solicitation* (*solícit*) also candidates for relative prominence preservation, and therefore its failure (1987a: 250). Such a proposal necessitates word-initial degenerate feet and rule ordering, and is therefore not pursued here.

and consistent success of foot-head preservation is assumed unquestioningly, e.g. Burzio (1994), Pater (1995, 2000), Benua (1997), Marvin (2002).

In spite of the consensus in favour of foot-head preservation, it is not totally uncontroversial. In the following sub-sections, apparent challenges to the status of foot-head preservation are considered. In §3.3.1, instances of apparent foot-head preservation failure are considered. In §3.3.2, a challenge from the opposite direction is considered: words which are not candidates for preservation but which nevertheless display the stress pattern that is supposed to be exclusively the result of foot-head preservation. Finally, in §3.3.3, Sainz's (1992) argument against foot-head preservation is considered.

3.3.1 Failure of foot-head preservation

As we saw in §3.2, examples from Kenyon & Knott (1944/53) have been used to reject the existence of relative prominence preservation. However, Kenyon & Knott (1944/1953: xxv) also give an example of variable secondary stress in a light-initial word that must be a candidate for foot-head preservation: ***à**cademícian* ~ *ac**à**demícian*. Similarly, Kager (1989: 170) gives ***à**ristocrátic* ~ *ar**ì**stocrátic*, and *c**à**nalizátion* ~ *can**à**lizátion*. Kager (1989: 170; see also Kager, 1995b: 20) argues that this variable secondary stress is caused by there being more than one plausibly embedded word, each with a different stress pattern – it is not evidence for the failure of foot-head preservation. Examples from Kager are given in (11):

(11) Variable stress in light-initial complex words (Kager, 1989: 170)

(a)	àc**a**demícian – àc**a**démic	cp.	(b)	acàdemícian – acádemy
	àr**i**stocrátic – ár**i**stocrat	cp.		arìstocrátic – arístocrat
	càn**a**lizátion – cán**a**lize	cp.		canàlizátion – canálize

The variable secondary stress in the embedding words in (11) therefore does not constitute evidence against foot-head preservation, but rather evidence for it.

There are, however, instances where foot-head preservation failure cannot be accounted for by there being more than one embedded word, such as Hayes' (1982) example *mìscegenátion~miscègenátion*. This word is only a candidate for foot-head preservation (#LLL pre-tonic sequence), and, according to Hayes, there is only one putatively embedded word: *miscégenate* (Hayes, 1982: 263). The initial-stressed variant is therefore an example of foot-head preservation failure. Hayes gets around this exception to foot-head preservation by proposing that the variable secondary stress is an example of inter-speaker variation, with some speakers having the word *miscégenate* and therefore *miscègenátion*, but other speakers lacking *miscégenate* and therefore having *mìscegenátion* with the default *àbracadábra* pattern. However, this is merely an assumption made by Hayes, and the possibility that foot-head preservation is less than successful within a single grammar must also be considered. As proposed for relative prominence preservation in §3.2.1, we have to consider the possibility that foot-head preservation is a variably, rather than one-hundred percent, successful phenomenon.

The examples in (11) also bring up another issue: which aspects of metrical structure are preserved? For the examples in (11a), Kager attributes the initial-syllable secondary stress to the preservation of weak positions – stresslessness – from the embedded word (shown in bold). Preservation of primary stress certainly receives the great bulk of the attention in the literature, but cases for the preservation of secondary stress have been made: Hammond (1989: 140-1) argues for secondary-stress preservation in the context *mànif**ĕ̀s**tátion* ← *mánif**ès**t* cp. *fòrestátion* ← *fórest*, and ***à**cademícian* ← ***à**cadémic* from (11a) could similarly be interpreted as such.[12] The issue of which aspects of metrical structure can be preserved has not been conclusively settled in literature since. In this thesis, I focus upon what can be construed as preservation of the embedded word's primary stress, but I will give preservation of other aspects of metrical structure some brief consideration in chapter 5.

3.3.2 Exceptional monomorphemic words

In (3), the generalisation given for #LLL words which are not candidates for preservation was that there should be consistent, initial-syllable secondary stress: ***à**bracadábra*. However, Bermúdez-Otero & McMahon (2006), Coleman (2000),

[12] Halle & Vergnaud (1987a) reject this being preservation of secondary stress and "maintain the more restrictive claim that only primary stresses are preserved cyclically" (Hammond, 1989: 141).

Halle & Kenstowicz (1991) and Hammond (1989) all point out that monomorphemic #LLL words do not always exhibit the ***à****bracadábra*-pattern: there are examples of #LLL monomorphemic words like *Ap****ò****llináris, ap****ò****theósis*, and *Ep****à****minóndas* with second-syllable secondary stress. These examples are problematic because they display the stress pattern that is supposed to be the exclusive preserve of second-syllable preservation in #LLL words. If purely phonological stress assignment can produce the second-stressed pattern, then we must question whether examples like *orìginálity* are examples of preservation, rather than purely phonological stress assignment.

Although the *Epàminóndas* problem absolutely should not be cast aside without proper empirical evaluation, past literature does suggest that these exceptional monomorphemic words may not be fatal to the argument for second-syllable foot-head preservation. The suggestion by both Bermúdez-Otero & McMahon (2006) and Halle & Kenstowicz (1991) is that the second-stressed pattern in #LLL monomorphemic words is exceptional: while there may be some overlap, it is not total and therefore not fatal. This is indeed the result I find in my extensive empirical investigation in chapter 5. Most importantly, Halle & Kenstowicz propose that *Epàminóndas*-type words may occupy a well-defined group of exceptions that may be accounted for accordingly:

> These words seem to be restricted to the Greek sector of the vocabulary and have another idiosyncrasy: they begin with a vowel. The simplest solution is to posit a special rule marking the initial, onsetless syllable extrametrical (Halle & Kenstowicz, 1991: 492).

Halle & Kenstowicz's etymological restriction is somewhat controversial – speakers may not necessarily be aware of such distinctions. However, the onsetless nature of the word-initial syllables is more promising, and is properly explored in chapter 5.

In sum, examples like *Epàminóndas*, whose stress patterns match that produceable by second-syllable preservation, add a second complication to the status of left-edge foot-head preservation, the first being the failure of foot-head preservation in *miscégenate* → *mìscegenation~miscègenátion*. However, neither complication is automatically damning to the argument for foot-head preservation; rather, both issues require further investigation.

3.3.3 Sainz's (1992) rejection of foot-head preservation

We saw in §3.2.1 that Sainz (1992) rejects relative prominence preservation; in this, Sainz was not alone. However, Sainz also takes the more controversial step of rejecting the existence of foot-head preservation.

Sainz (1992) proposes that the highest stratum in Lexical Phonology has noncyclic stress assignment. Sainz's model therefore somewhat resembles the stratal model of phonology proposed by Halle & Vergnaud (1987a) and Halle & Kenstowicz (1991), where the default formulation of the first stratum was noncyclicity (see §1.2.2.1). The crucial difference between Sainz's model and these is that Sainz denies that a weak stress preservation mechanism is necessary at all, and so does not introduce any ad hoc devices like Stress Copy to deal with it.

Sainz presents evidence which is argued to indicate that the first stratum has no stress cycle: (i) the putative failure of relative prominence preservation discussed in §3.2.1; (ii) allegedly incorrect results produced by particular orderings of cyclic stress rules in Kiparsky (1982) (**expèctátion*, not *èxpĕctátion*) (Sainz, 1992: 101); and (iii) the inconsistency of putative pre-tonic stress preservation (Sainz, 1992: 119-20). I take none of these arguments to be damning evidence against stratum one stress preservation. With respect to the failure of relative prominence preservation, I showed in §3.2 that this data requires further investigation. Second, the problems of cyclicity and rule ordering are irrelevant here, as I will be analysing weak stress preservation in Stratal OT, a model which rejects rules altogether. Finally, we saw in chapter 2 that the inconsistency of pre-tonic preservation (of vowel-quality, not stress, in our argument) was not evidence against its very existence; rather, it appears to be a probabilistic phenomenon, as I suggested in §3.2 may also be the case for relative prominence preservation.

With respect to foot-head preservation, Sainz proposes that the locations of foot heads in complex and simplex words alike can be determined by the contents of the underlying representation alone; no device like the cycle is required to preserve aspects of surface structure like stress. Sainz's rejection of cyclic foot-head preservation crucially relies upon her assumption of an underlying three-way vowel distinction: long, short and reduced. Sainz argues that non-alternating schwas are underlying and unstressable, and that this can account for the secondary stress placement in words which are traditionally argued to have their secondary stress assigned by the cycle: the stress contours of complex words do not "follow from cyclic application of stress rules, but instead from purely phonological properties of the underlying forms of those words" (Sainz, 1992: 138).

Sainz's analysis absolutely denies that complex words inherit surface phonology like stress from their bases. For example, there is second-syllable stress in *orìginátion* not because it inherits stress from *oríginate*, but because its underlying representation is argued to be /ɒ.**rì**.dʒɪ.néɪ.ʃən/ and so her relevant right-to-left stress rule assigns stress to the second syllable (Sainz, 1992: 137). In contrast, Sainz argues that there is initial-syllable stress in *pèregrinátion* because its underlying form is /pɛ̀.**rə**.grɪ.néɪ.ʃən/ – the right-to-left stress rule is blocked from assigning stress to the second syllable, and so it is forced onto the initial syllable.

A potential problem with Sainz's analysis is that, at least in Jones (2003) and Wells (2000), there are examples of putatively initial-preserving words which do not have non-alternating schwa in the second syllable, e.g. *èd*[ɪ]*ficátion* ← *édify*. Although [ɪ] is often in free variation with schwa in English (Giegerich, 1992) and so could be said to function as a reduced vowel of sorts, it can also occur in necessarily stressed syllables, e.g. *bit*. There is therefore nothing to block **edìficátion* in Sainz's analysis. As it stands, Sainz's particular analysis does not constitute an argument against left-edge foot-head preservation.

3.4 Testing for left-edge stress preservation

In this section, the observations from §3.1-§3.3 are brought together in order to establish which syllable weight environments will yield evidence for or against left-edge stress preservation.

3.4.1 Testing for foot-head preservation: #LLL, #LLH and #LHL words

In some instances, the predictions of purely phonological stress assignment and those of foot-head preservation overlap. Because more than one mechanism could be assigning stress in these contexts, they cannot be used to provide evidence for stress preservation.

With respect to second-syllable foot-head preservation, the #LHL context is unlikely to provide any evidence for preservation. We saw in (3) that secondary stress is predicted to fall on the heavy second syllable of this sequence in words which are not candidates for preservation, e.g. *Mon**òn**gahéla*. Consequently, even though it may be argued that stress preservation accounts for the second syllable stress of *ad**àp**tabílity* (← *ad**áp**table*), there is no evidence for this: purely phonological, weight-sensitive stress assignment could also account for the secondary-stress placement in *adàptabílity*. The #LLL context is more promising, as the predictions of foot-head preservation and pure phonology should conflict in this context. Although we have acknowledged that #LLL monomorphemic words like *Epàminóndas* do occur, the initial-stressed pattern seen in *àbracadábra* is also possible, whereas it should not be with second-syllable preservation – *oríginal* → **òriginálity*.

The evidence for initial-syllable foot-head preservation, e.g. ***vè**rificátion* (← ***vé**rify*), is likely to be much less strong. We may find some contrast in the #LLL pre-tonic context: although exceptions to the *àbracadábra* stress pattern, like *Epàminóndas*, are potentially problematic for second-syllable preservation, such exceptions do contrast with the predictions of initial-syllable preservation. However, by virtue of the fact that second-stressed examples like *Epàminóndas* are argued to be exceptional, the contrast may be minimal. A context where we could expect the contrast to be much greater is where foot-heads are routinely predicted to fall on the second syllable at the expense of the first in monomorphemic words: #LHL. The predicted problem here is that, for independent reasons, it is unlikely that initial-syllable preservation would arise in the #LHL context: English primary-stress assignment does not skip over a heavy syllable to assign stress to a preceding light syllable,[13] meaning there is no way the initial syllable could get stress to be subsequently preserved: *****é**ssential* → *****è**ssentiálity* cp. *ess**én**tial* → *ess**èn**tiálity*.[14] This hypothesis is supported by the results presented in §5.4.

---

[13] Embedded words are generally trisyllabic or longer, ruling out the possibility of a following heavy syllable being extrametrical in the original stress assignment to the embedded word.

[14] Apparent exceptions to weight-sensitive main stress assignment such as ***cý**linder* and ***bá**dminton* are potentially otherwise explained: Giegerich (1999: 18) proposes non-syllabicity of the final /r/ at the

A final context to consider for both initial and second-syllable foot-head preservation is #LLH. This context receives less attention in the literature, but there is no reason why we should not consider it also. In monomorphemic words, an initial-stressed pattern is expected, e.g. *H**à**licarnássus*. With second-syllable foot-head preservation, this may be overridden, e.g. *dev**è**lopméntal* (*devélop*). With initial-syllable preservation, e.g. ***à**limentátion* (*áliment*), the predictions of pure phonology and foot-head preservation completely overlap. A fourth, possible pre-tonic syllable-weight context, #LHH, is not considered – this combination is not discussed in monomorphemic words in the literature (§3.1, above).

The phonological contexts where the predictions of left-edge foot-head preservation and purely phonological stress-assignment conflict are summarised in table 2. This table and the next are indebted, in part, to an original table by Hammond (1989: 145) and a reproduction by Halle & Kenstowicz (1991: 490). (Numbers are used to make the stress patterns explicit: 1= main prominence, 2 = secondary prominence, and so on.)

---

time stress is assigned to account for *cýlinder*, and Burzio (1994) argues that syllables closed by sonorants may optionally function as light, potentially accounting for *bádminton*.

	#LLL	#LHL	#LLH
Monomorphemic	**à**bracadábra (2-0-0-1) Ep**à**minóndas (0-2-0-1)	Mon**òn**gehéla (0-2-0-1)	H**à**licarnássus (2-0-?-1)[15]
Initial-syllable preservation	v**è**rificátion (2-0-0-1)	*not anticipated*	**à**limentátion (2-0-?-1)
Second-syllable preservation	or**ì**ginálity (0-2-0-1)	ad**àp**tabílity (0-2-0-1)	dev**è**lopméntal (0-2-0-1)

Table 2: predictions of foot-head preservation versus purely phonological stress-assignment

3.4.2 Testing for relative prominence preservation: #HHL and #HLL

Two important issues must be noted with respect to initial-syllable preservation in #HHL and #HLL words. First, a crucial point that has not been made explicit so far in this chapter is that relative prominence preservation is only likely to be relevant to **second**-syllable preservation. This is because relative prominence preservation is only relevant if both the initial and second syllables of the embedded word are stressed, as in *sènsátional*. In heavy-initial embedding words with initial main stress, e.g. *óxygenate*, the second syllable is not stressed precisely because this syllable has been skipped over for main stress. There is therefore no relative prominence of stresses to preserve; rather, this is another instance of foot-head preservation. Following on from this, we do not expect examples of initial-syllable foot-head preservation in #HHL words, for the same reason that we do not expect initial-syllable preservation with #LHL words: with a heavy second syllable, primary stress

[15] It has been proposed that the pre-tonic syllables of #LLH words are sometimes stressed (Halle & Kenstowicz, 1991: 492, f.n. 7). This argument may need to be reassessed in light of the rejection of pre-tonic stress preservation in chapter 2, although the argument made there is not that pre-tonic syllables **cannot** be stressed, but that, if the absence of vowel reduction is used as the only evidence for the presence of such stress, then the theory of word stress becomes problematic.

is unlikely to ever be assigned to the initial syllable in the embedded word.[16] (This hypothesis is supported by the data in §4.5.1.)

The predictions for #HLL and #HHL words are summarised in table 3. There are no predictions for two other possible heavy-initial combinations, #HHH and #HLH, as examples of monomorphemic words with these pre-tonic syllable weight combinations are not discussed in the literature (§3.1, above).[17]

	#HLL	#HHL
Monomorphemic	L**ù**xipalílla (2-0-0-1) D**ò**d**ĕ**canésian~ D**ŏ**d**è**canésian (2-3-0-1 or 3-2-0-1)	Tìc**ŏn**deróga~ Tīc**òn**deróga (2-3-0-1 or 3-2-0-1)
Initial-syllable preservation	**òx**ygenátion (2-0-0-1)	*not anticipated*
Second-syllable prominence preservation	**ăntì**cipátion (3-2-0-1)	s**ŏm**n**àm**bulátion (3-2-0-1)

Table 3: predictions of left-edge preservation versus purely phonological stress-assignment in #HLL and #HHL words

Again, it is clear that the predictions of purely phonological stress-assignment and preservation sometimes overlap: for example, in #HLL words, initial-syllable preservation and purely phonological stress-assignment may well produce the same result: *òxygenátion* cp. *Lùxipalílla*. However, the potential existence of exceptions like *Dòdĕcanésian~Dŏdècanésian* means that there may nevertheless be marginal

[16] Weak preservation occurs in stratum one vocabulary, and therefore any embedded words should obey the Latinate main stress rule. This means that, even if an embedded word is itself complex, the embedded word's stress should obey the rules for main stress assignment in monomorphemic words.
[17] Possibe monomorphemic and complex examples have been found in my empirical research: see chapter 4.

evidence for initial-syllable preservation in #HLL words – the contents of each row are again different.

Finally, it is worth noting that, even if relative prominence preservation fails, #HLL complex words may still provide evidence for second-syllable foot-head preservation. This is shown in table 4:

	#HLL	#HHL
Monomorphemic	Lùxipalílla (2-0-0-1) Dòdĕcanésian~ Dŏdècanésian (2-3-0-1 or 3-2-0-1)	Tìcŏnderóga~ Tĭcònderóga (2-3-0-1 or 3-2-0-1)
Second-syllable foot-head preservation	ăntìcipátion~ àntĭcipátion (3-2-0-1 or 2-3-0-1)	sŏmnàmbulátion~ sòmnămbulátion (3-2-0-1 or 2-3-0-1)

Table 4: predictions for foot-head preservation in heavy-initial words

Failure of relative prominence preservation would predict a complete overlap in the behaviour of #HHL complex words where second-syllable preservation can occur and #HHL monomorphemic words, shown in the column on the right in table 4. But, as #HLL words are not generally expected to have their second syllable stressed on purely phonological grounds, we may still have evidence for second-syllable foot-head preservation in #HLL words, as the difference between the two rows in table 4 shows.

3.5 Conclusion

In this chapter, I have explored the traditional arguments surrounding the weak stress preservation which occurs at the left edges of words with three pre-tonic syllables:

left-edge stress preservation. There have been arguments for two distinct sub-types of left-edge preservation: relative prominence preservation, applicable to words with a heavy initial syllable, and foot-head preservation, applicable to words with a light initial syllable. Arguments have been made which question the status of both types of left-edge stress preservation.

With respect to relative prominence preservation, Kiparsky (1979) proposed that the relative prominence of feet is preserved under embedding: *sèns**á**tional* → *sĕns**à**tionálity*. Relative prominence preservation is logically impossible in phonotactically parallel monomorphemic words, hence variation in secondary stress placement is anticipated: *T**ì**cŏnderóga* ~ *Tĭc**òn**deróga*. The existence of relative prominence preservation has been argued against in light of examples where relative prominence preservation fails: e.g. *ìc**ó**noclast* → ***ì**cŏnoclástic~ĭc**ò**noclástic*. I argued that the failure of relative prominence preservation in some words does not automatically imply the complete absence of relative prominence preservation – relative prominence preservation may be a variably successful phenomenon.

The argument against relative prominence preservation does not preclude the existence of foot-head preservation: foot-heads may be preserved, if not their relative prominence contours. The existence of foot-head preservation has been generally less controversial in the literature, with the exception of Sainz (1992). However, I identified two complications to showing the existence of foot-head preservation: (i) failure of foot-head preservation in *misc**é**genate* → *m**ì**scegenátion~misc**è**genátion*; (ii) the fact that some monomorphemic words appear to have the same pattern as is expected of second-syllable foot-head preservation: *Ep**à**minóndas* ≈ *or**ì**ginálity* (← *or**í**ginal*). Both issues mean that thorough empirical research into foot-head preservation is required if the phenomenon's existence is to be assured, and its precise nature understood.

In conclusion, the empirical status of left-edge stress preservation is still controversial, and further empirical investigation of the phenomenon is clearly required. This is precisely the conclusion arrived at by Halle & Vergnaud (1987a):

> [F]or words like *solicitation* the alternative with secondary stress on the first syllable seems to be excluded for everyone we have consulted. The obvious factor at work here is etymology: since *solícit* has stress on the second

> syllable, this stress is preserved in the derived noun. This cannot be the whole story, however, (as shown by many of the words cited in footnote 12 [e.g. *academician*, *acceptability*, *deceleration*]) the etymology is often not decisive. We have not found an especially insightful solution to this problem and **leave it with the observation that the topic requires further study** (Halle & Vergnaud, 1987a: 250; boldface SC).

The following chapters take up Halle & Vergnaud's suggestion.

In light of the arguments reviewed in this chapter, any empirical investigation would be expected to come to one of the following four conclusions:

(i) There is no weak stress preservation of any kind (Sainz, 1992).

(ii) There is foot-head preservation, but no relative prominence preservation (Liberman & Prince, 1977; Halle & Kenstowicz, 1991; Burzio, 1994).

(iii) There is both foot-head and relative prominence preservation (Kiparsky, 1979).

(iv) There is both foot-head and relative prominence preservation, but both achieve variable success – the new hypothesis formulated here.

In chapters 4 and 5, I explore data from two pronouncing dictionaries of English, Jones (2003) and Wells (2000), in order to discern the status of left-edge stress preservation. I obtain evidence for preservation by comparing the stress behaviour of complex words where preservation is anticipated to phonotactically parallel words which are not candidates for preservation, as outlined in §3.4. There we will see that hypothesis (iv) is supported.

Part II: empirical investigation

Chapter four: the data for left-edge stress preservation I: heavy-initial words

4.0 Introduction

In chapter 3, it was shown that there is controversy over the status of left-edge stress preservation. A thorough investigation into the data for left-edge stress preservation is therefore carried out over the next two chapters. In this chapter, words with three pre-tonic syllables where the first syllable is heavy are considered. It is shown that there is evidence for both foot-head and relative prominence preservation in words with heavy initial syllables. However, it will be seen that relative prominence preservation in these words is not uniformly successful, but rather appears to be a variably successful phenomenon. Words with light initial syllables are treated in chapter 5.

The outline of this chapter is as follows. In §4.1, the methods used in collecting and categorising the data are outlined. In §4.2, the data for words which are candidates for relative prominence preservation and second-syllable foot-head preservation are presented. These data consist of words whose embedded words have main stress on their second syllable, e.g. *anticipation* (*àntícipate*) and *somnambulation* (*sòmnámbulate*). Phonotactically similar words which are not candidates for stress preservation are presented in §4.3, and comparisons with these words are used to generate evidence for relative prominence and second-syllable foot-head preservation in §4.4. In §4.5, the data and evidence for initial-syllable preservation in heavy-initial words – words like *gentrification* (*géntrify*) – are presented. The chapter concludes in §4.6.

All percentages given in this chapter and the next are approximated to 1 decimal place.

4.1 Method

The data examined in this chapter and the next have been collected from two English pronouncing dictionaries: Jones (2003) and Wells (2000). In this section, the nature of the data is discussed, and the decision to use pronouncing dictionary data is justified. The way in which the data has been categorised is then outlined.

4.1.1 The pronouncing dictionary data

The aim has been to collect an exhaustive list of the five- and six-syllable words with three pre-tonic syllables that are in two English pronouncing dictionaries: Jones (2003) and Wells (2000). The data has been collected by hand from the printed versions of the dictionaries, but effort has been made to try and ensure that the lists are complete. The pronouncing dictionary data presented here constitutes the most careful examination of left-edge stress preservation data to date.

Jones (2003) and Wells (2000) are pronouncing dictionaries each of approximately 80,000 entries; both dictionaries give preferred and alternative pronunciations, American English pronunciations and, in the case of Wells (2000), British regional pronunciations – Jones (2003) gives regional pronunciations for place names only. The main pronunciations given by Wells (2000) correspond to a "broader RP" model (Wells, 2000: xiii). Jones (2003: v) chooses to abandon the term 'RP' altogether in light of its connotations of high social class, instead describing its pronunciation model as the BBC English "of broadcasters with an English accent". It is, therefore, fair to say that, while the main pronunciations given by both dictionaries try to avoid regional pronunciations, the accent is perhaps not as highly elitist as that modelled in earlier pronouncing dictionaries. However, as noted below, even if the accent of the dictionary's pronunciation model was highly elitist, this would be unproblematic for the present study.

Data from pronouncing dictionaries provides the basis of the discussions of left-edge stress preservation in Halle & Vergnaud (1987a), Sainz (1988, 1992) and Kager (1989) (§3.2), as well as of more recent work on putative pre-tonic stress preservation by Hammond (2003a, b) (§2.5.1). However, simply repeating methodology was not the main reason dictionaries were chosen; rather, dictionaries were felt to provide the most suitable data for the study here. There are databases of words that are computer-searchable by stress pattern: the *MRC Psycholinguistic Database* (Wilson, 1987); Michael Hammond's own *English on-line dictionary* (Hammond, 2005); and CELEX (Baayen et al., 1995). Using one of these databases would have been less laborious, but all fell short on at least one of the following three criteria:[1]

(i) Providing some definitive information about the originating sources of the data contained in the database – particularly problematic for Hammond's *English on-line dictionary*.

(ii) Providing detailed information with respect to alternative stress patterns and segmental realisations – particularly problematic for the *MRC Psycholinguistic Database*.

(iii) Being the most recent editions: the English version of CELEX is based on pronouncing dictionaries from the 1970s.

In contrast, both Jones (2003) and Wells (2003) fulfil all of criteria (i) to (iii). It was also important to ensure that the data did not contain inter-dialectal variation: the data needs to resemble an idiolect as far as possible to be in accordance with Generative Grammar's 'ideal speaker-hearer'. Generative linguistics aims to study 'internalised language', or 'I-language' – "some element of the mind of the person who knows the language, acquired by the learner, and used by the speaker-hearer" (Chomsky, 1986: 22). Each speaker of a language may have a different I-language, each one thus corresponding to a different grammar. To avoid mistaking variation between grammars for variation in a single grammar, it is vital that just a single idiolect is studied. Corpuses of recorded speech were, therefore, also rejected as a source of data. In contrast, pronunciations indicated to be those of American English or regional British English can be easily factored out of the pronouncing dictionary data.

It is important to note that the common criticisms of the prescriptivism of pronouncing dictionaries do not impinge upon the present research interests, nor does the fact that dictionary pronunciations are an abstracted form of data. Hartman (2000: 250) argues that "prescription must be based on something": pronouncing dictionaries still contain potentially useful observations about language, regardless of their social implications. For each reference accent, the dictionaries aim to create a pronunciation model that consists of a single homogenous idiolect. This model is

[1] In any case, Jones (2003) is available with a CD-Rom that is searchable by stress pattern. This was used in the present investigation to double-check some of the results.

based upon solid real-world sources: the intuitions of pronunciation editors who are themselves experts in their field, and fellow speakers of English they have consulted or observed. That pronouncing dictionaries contain generalised pronunciations, with much regional and social variation necessarily factored out (Hartman, 2000: 252), is preferable for the present research: the only stress variation which is of interest here is that which occurs in the absence of inter-dialectal or social variation, and a pronouncing dictionary offers precisely such moderated data.

However, there are limitations on the claims which can be made on the basis of this investigation of pronouncing dictionary data. First, it cannot be claimed that the artificial idiolects of the pronunciation models of Jones (2003) and Wells (2003) definitely correspond to a single real-world speaker of English; it may not, therefore, be the case that stress variation recorded in the dictionary is definitely real-world intra-speaker variation. This is apparently problematic with respect to Chomsky's concept of a single ideal speaker-hearer: we cannot be said to be modelling the linguistic knowledge of a single individual. However, when one considers that the ideal speaker-hearer is itself an abstract theoretical concept, the data being proposed here – artificial, abstracted and internally coherent idiolects – seems altogether more reasonable. (It is also important to note that what is being modelled here cannot be said to be part of the stress system 'of English'. Two artificial idiolects of English speakers are being modelled.)

A final likely point of contention with respect to the use of pronouncing dictionaries is that many of the words which they contain are not known to the average English speaker. This issue is particularly relevant here, as the data used in discussions of left-edge preservation consists of long, often rather obscure words.[2] However, this opposition to the use of pronouncing dictionary data does not stand for

---

[2] Only the long loan words considered here can reveal the phonological generalisations which would remain hidden if short, morphologically simple words were considered. Weak stress preservation can logically only occur in complex words, and so it is likely to occur in longer words; however, short, monosyllabic words constitute the largest part of day-to-day English vocabulary (Teschner & Whitley, 2004: 19, 27). Additionally, the vocabulary in which weak stress preservation occurs is likely to be specialised or prestigious. Weak stress preservation is associated with stress-shifting affixation, and stress-shifting affixes are generally Latinate; in turn, Latinate suffixes prefer to attach Latinate bases (the [± Latinate] constraint; Saciuk, 1969; Aronoff, 1976; Booij, 1977; Giegerich, 1999), and Latinate vocabulary tends to be more specialised or prestigious. A similar situation exists in German: Alber (1998) has to restrict her discussion of German stress preservation to "loan" words, native Germanic vocabulary being mostly monosyllabic. As long as a speaker has even a few of the words considered in this thesis, with their preserving stress patterns, then there is something to explain.

two reasons: (i) I do not claim to be modelling the grammar of all speakers of English, but rather of the artificial idiolect of a pronouncing dictionary; (ii) the words in question are known to the speakers of English who have advised upon the dictionary pronunciations.

It is worth pointing out the different histories of the two pronouncing dictionaries in relation to the stress pronunciations they give (this issue becomes particularly relevant in §6.5.2). Jones (2003) is the latest edition of Daniel Jones' original 1917 pronouncing dictionary. A current editor of Jones' dictionary, Peter Roach (personal communication), advises that Daniel Jones did not use any systematic rules in assigning stress to polysyllabic words when compiling the original dictionary, and that subsequent editors have revised individual pronunciations only as felt to be necessary. Wells (2000) has a much shorter history, being the latest edition of John Wells' original 1990 pronouncing dictionary. John Wells (personal communication) advises that, while systematic rules have not been used to decide the stress pronunciations he gives in Wells (2000), nor are the recommended pronunciations entirely intuitive; rather, phonological factors, such as a word's intonation behaviour, have been used for guidance.

4.1.1.1 Divergence from dictionary conventions

The most important change which has been made with respect to the dictionary conventions concerns the Maximal Onset Principle (MOP; Selkirk, 1982b: 359). The dictionaries show stressed syllables attracting onsets from following syllables, but resyllabification in contravention to the MOP is generally argued to be a triggered by stress assignment to a light syllable, rather than the cause of it. To distinguish between syllable weight predictable from the phonemic strings of underlying representations, and weight adjustments forced at a later stage by stress assignment to a light syllable, the MOP has been strongly respected in the categorisation of the data. All word-internal clusters that could be syllabified as a complex onset, including /sC/ clusters, have been syllabified as onsets in the first instance; the controversial nature of this step with respect to /sC/ clusters is taken into account in the subsequent data analysis where possible. It is assumed that prefixes syllabify with the following word if this is phonotactically possible, unless a dictionary shows a prefix as consistently syllabifying separately for all stress variants of a word, e.g. [dɪs.sæ] is assumed for *dissatisfaction* [dɪ̀s.sæ]~[dɪs.sæ̀] (Jones, 2003), not [dɪ.sæ].

Two changes have been made to the segmental representations given in the dictionaries. Where the dictionaries use the symbol /e/ to denote a lax vowel, it has been replaced with /ɛ/, following the notation of Giegerich (1992). The short, tense, unstressable vowels described as 'weak' by the dictionaries – e.g. [i] in *happ**y***, [u] in *sit**u**ation* (cp. [iː, uː]) – have been grouped in with short, lax [ɪ, ʊ] here, in line with more traditional transcriptions. This does not represent a disagreement with the pronunciations proposed in the dictionaries, but avoids complicating the vowel system any more than necessary for the present research by retaining a systematic correlation between tenseness and length.

4.1.2 Categorising the data

In order to categorise the data taken from the two pronouncing dictionaries with respect to stress preservation, decisions had to be made as to how to interpret the data with respect to two factors: stress, and morphological complexity.

4.1.2.1 Stress

Three issues came up in the categorisation of word stress patterns. First, the dictionaries only recognise primary and secondary levels of stress, so that any argument for a tertiary level of prominence is theoretically motivated and is not confirmed by the dictionary contents: a word like *ambassadorial*, is represented as ***am**bàssadórial*, not ***ăm**bàssadórial*. Second, the dictionary entries are themselves citation forms (Wells, 2000: 143): by definition, citation forms are not examples of emphatic stress or connected speech stress-shift triggered by adjacent words. Kenyon & Knott (1944/53) do suggest that there may be difficulty in isolating a 'citation' stress pattern for certain words (§3.2.1), and just how citation-like Jones' and Wells' dictionary entries are is an issue addressed later in §6.5.4.2. Third, words with a main and alternative stress pattern have been treated as variation within the artificial idiolects of the dictionaries' pronunciation models. This last point requires some elaboration.

Words have been classed as examples of variable secondary stress only if the same segmental representation for the first and second syllables occurs with both stress patterns. For example, [bə̀ʊ.lɪ]*viano* and [bɒ.lɪ̂]*viáno* for *Boliviano* would not be classed as variable secondary stress, but rather as two separate forms from two different underlying representations (for this reason, some words may appear in the investigation more than once, as each underlying representation is treated as a separate lexical item). An exception to this point is made for the alternation of a full vowel and schwa: vowel reduction is generally accepted to be a consequence of stresslessness, rather than the cause of it, vowel reduction applying after all other stress rules (e.g. Halle & Vergnaud, 1987a: 240);[3] a shared underlying representation is therefore easily and reasonably discerned.

Stressed prefixes confuse the identification of variable stress in words where second-syllable preservation is anticipated. Any instances where the dictionary shows prefixes to be stressed as well as the second syllable of the word, e.g. ***ìnè**dibílity*, have not been treated as examples of variable relative prominence: it could be argued that the prefix has received stress in its own right upon productive prefixation with *in*- of *èdibílity*.[4] Wells and Jones both indicate clear variation in prominence between prefix and second syllable in other instances, as in ***ìn**finitíval~in**fi**nitíval*, and so the stress shown by Wells on the prefix in ***ìnè**dibílity* must be assumed to be empirically different to the stress which is not shown on the prefix by Wells in ***in**finitíval*, but is shown in ***ìn**finitíval*. Words which are represented as in ***ìn**finitíval~in**fi**nitíval* (*in**fí**nitive*) alone are treated as instances of failure to consistently preserve second-syllable prominence.[5]

---

[3] Of course, I argue in chapter 2 that it is the absence of stress **plus** other factors which brings about vowel reduction. The argument I am making here is crucially concerned with the ordering of stress and vowel reduction.

[4] Hurrell (2001) proposes a stratum two rule of 'prefix-footing'; analyses which admit prosodic words (e.g. Booij & Rubach, 1984; Spzyra, 1989; Raffelsiefen, 1999) propose that productive prefixes may be stressed as a consequence of being a separate prosodic word.

Although stressed prefixes are generally only argued to occur in words which could be the result of productive prefixation (Hurrell, 2001), the dictionaries do give pronunciations like *d**è**c**è**lerátion*. The root of *deceleration* [-*celer*-]$_R$ is clearly bound, precluding productive prefixation. Theoretical arguments and dictionary data are therefore not completely congruent, but this does not negate the argument for excluding 'both-stressed' forms like *d**è**c**è**lerátion* on the grounds that they may be empirically different from alternating forms like ***ìn**finitíval~in**fi**nitíval*.

[5] Words for which variable and both-stressed pronunciations are given – *d**è**activátion~de**àc**tivátion* and *d**èà**ctivátion* – are included on the merit of the former variable pronunciation only. Words are not excluded if their embedded words have prefix stress, e.g. *d**è**áctivate* embedded in

A final point with respect to the categorisation of stress patterns is the status of non-alternating schwa. Sainz (1992) proposed that non-alternating schwa is underlying: a three-way distinction between underlying vowels (tense, lax, reduced) is vital to Sainz's rejection of the stress cycle (§3.3.3) in favour of the argument that syllables with underlying schwa reject stress. To have an underlying schwa is reasonable given standard generative constraints upon abstractness for non-alternating morphophonemes, but underlying schwa is not generally exploited in stress literature as the **cause** of failure to receive stress, cp. Sainz's proposal. Words with non-alternating schwas have not been excluded from analysis altogether in the following chapters, but non-alternating and alternating schwas are distinguished between, and forms with non-alternating schwa are excluded from analyses where possible. The vowel [ɪ] is often in free variation with [ə] as a reduced vowel in English, but as [ɪ] also occurs in stressed syllables (Giegerich, 1992: 285), [ɪ] has been treated as stressable.

4.1.2.2 Morphological complexity

A crucial part of categorising the data was determining from which word stress is potentially preserved. In some instances, a word can be argued to have more than one possible embedded word: e.g., for *demòdulátion*, either *demódulate* or *mòdulátion* (supporting discussions of *de*- and -*ation* are given in Marchand (1969: 156, 261)). As long as there was a primary stress on the first or second syllable in at least one of the potentially embedded words, the embedding word has been included as a candidate for left-edge stress preservation in this chapter, but the possibility of secondary stress preservation will not be ignored altogether (see §5.5.1). In the case of obscure words which could be naively perceived as monomorphemic or being built directly on a bound root, the following steps were taken to decide whether or not they were, in fact, candidates for preservation:

(i) If a morphophonologically similar, shorter word which could plausibly be embedded was in the same pronouncing dictionary, and OED definitions (consulted when I was unfamiliar with one or both words)

dèactivation~deàctivátion, as a difference in prominence is clearly apparent in ***dèáctivate*** which may be expected to be preserved in *deactivation*.

indicated that the two words were also semantically similar, a word was treated as complex.[6]

(ii) Where the pronouncing dictionary did not contain a plausibly embedded word, the OED was checked for any likely candidate (e.g. for *dilatability* from Jones (2003), *dilatable* was found in the OED).

(iii) Where a plausibly embedded word was in the OED (the OED being consulted under (ii)) but was marked as 'rare', 'obsolete' or not naturalised, it was disregarded, and the putatively embedding word was put in the monomorphemic and bound-root base category (e.g. †*pusillanime* for *pusillanimity* from Wells (2000)).

With respect to (ii), most often there was no candidate for an embedded word in the OED, and the potential embedding word was classed as bound-root base or monomorphemic. In the few instances where the OED contained a plausibly embedded word, the word is included, but the OED source is indicated on the accompanying appendices: the OED gives less exhaustive pronunciations, and words from the OED are outwith the artificial idiolect offered by a single pronouncing dictionary. It was nevertheless felt to be useful to consult the OED, as in some instances there were what were essentially accidental gaps in the dictionaries: not particularly obscure putatively embedded words were occasionally omitted from a dictionary, and not to include these words would present a misrepresentative picture of the language and would exclude much-needed data.

The issue of morphological complexity is also complicated by deciding which words should be allowed to be embedded. I consider preservation to be a possibility between any two semantically and morphophonologically similar words where one is longer than the other, and the shorter word has stress on the first or second syllable such that it could be expected to be preserved in the longer word. This decision means that, sometimes, the embedded word is truncated in the embedding word, e.g. *Antipodean* (*Antípode**s***), *comparability* (*compárab**le***),

[6] All references to the OED are to the *OED Online*, www.oed.com (January 2005-May 2006). Semantic similarity has been a judgement call: no strict parameters with respect to the number of shared meanings have been employed for difficult cases.

dissatisfaction (*dissátisf**y***), *oesophageal* (*oesóphag**us***).[7] My stance is implicit in some treatments of stress preservation: for example, Burzio (1994: 186) argues for preservation from *tot**á**lit**y*** in *tot**à**litárian*;[8] Kager (1995b) similarly argues for preservation from *demónstrab**le*** in *demònstrabílity*. Proposing that stress may be preserved from truncated words is not uncontroversial: Kiparsky (1982), and taking his lead Giegerich (1999: 24, 29), reject synchronic truncation on stratum one in Lexical Phonology on the grounds of formal cost (cp. Aronoff, 1976). In the absence of synchronic truncation, words embedding truncated words, just like words directly formed upon bound or cranberry roots, should be treated as monomorphemes and therefore classed as ineligible for stress preservation. It seems implausible, however, that speakers would never associate pairs of stratum one forms like *dissatisfaction* and *dissátisfy*.

The broad delimitation of the set of words which may be considered as candidates for left-edge stress preservation means that instances where the putatively embedding and embedded words are not separated by a recognised suffix of English (defined here as a suffix listed in Marchand (1969)) are also considered as candidates for preservation. Theoretical discussions of stress preservation tend to deal only with recognised suffixes, but a fully comprehensive survey like the present one cannot afford the luxury of such selectiveness. The absence of a salient suffix as a terminal element is not necessarily a bar to the perception of morphological complexity: Raffelsiefen (1993: 11-12) argues that speakers clearly recognise *hate* within *hatred*, yet, in an OED search, I have found only two other English words with the same *-red* suffix: *kindred*, and the altogether unfamiliar *gossipred*. Speakers have very little evidence, therefore, from which they can discern *-red* is a suffix of English, yet this does not prevent them from perceiving *hate* within *hatred*.

Some examples of putatively embedding words from the data set with non-suffixal terminal elements are given in (1). (The grouping of terminal elements in (1) is purely orthographic.) These words are exhaustively identified on the accompanying appendices.

---

[7] There are rare instances of very borderline cases, e.g. *semasiology*. *Semasiology* is clearly similar in meaning to *semantic*, but *semantic* would have to be very heavily truncated to preceding its second syllable coda in order to be embedded in *semasiology*. I made the judgement call to treat such examples as part of the monomorphemic and bound-root base category.
[8] See also Burzio (1994: 136) for examples of truncation.

(1) Non-standard terminal elements found in the complex word data

-(o)sis:	anastomosis (anástomose); Winnipegosis (Wínnipeg)
-ola:	Hispaniola (Hispánic)
-(a)ndo:	accelerando (accélerate)
-o:	inamorato (inámorate); ejaculatio (ejáculate)
-us:	tyrannosaurus (tyránnosaur)
-ano:	Boliviano (Bolívia)
-a:	impedimenta (impédiment); desiderata (desíderate),
-ology:	phenomenology (phenómenon); Assyriology (Assýria)
-rama:	Bananarama (banána)
-iasis:	elephantiasis (élephant)
-masia:	paronomasia (páronym)

In light of the parameters by which preservation candidates are determined, the category of words which are not candidates for preservation is defined as consisting of monomorphemes, and words formed directly upon (synchronically) bound (including cranberry) roots (e.g. *egalitarian*, *icosahedron*), where there is no shorter free form that could be embedded. Complex words that are constructed directly upon a bound root are not considered to be candidates for preservation, as stress can only be preserved from words with their own independent stress patterns, and therefore bound roots cannot logically be candidates for stress: they are never used by speakers in isolation.

4.2 Words where second-syllable preservation is expected

In chapter three, we saw that heavy-initial words where second-syllable preservation is anticipated are candidates for both relative prominence and foot-head preservation. Here, the data for this group of words are presented.

4.2.1 The data

The data from both dictionaries consist of words where the second syllable is shown to bear primary stress in the embedded word and the initial syllable is heavy, e.g. *somnambulation* (*somn**ám**bulate*). Out of the four possible trisyllabic syllable-weight sequences beginning with heavy syllables (#HHL, #HHH, #HLH, #HLL), the data appear to fit into just two of the sequences: #HHL and #HLL (the reader can verify this for themselves by consulting appendices A and B). There are no incontrovertible examples of the other two sequences, #HHH and #HLH.[9] This consistency may be at least, in part, the result of primary stress application to the embedded word. The majority of the embedded words are quadrisyllabic, and must therefore have antepenultimate primary stress in order to predict second-syllable preservation in the embedding word. As English main stress is weight-sensitive, it follows that these words cannot have heavy third (penult) syllables: this third syllable is not extrametrical, and so would receive main stress if it was heavy.

There was evidence of relative prominence preservation failure from Wells (2000). Excluding words where the first syllable contains a consistently reduced vowel and where the same segmental string did not occur with both stress pattern variants, a total of 153 #HHL and #HLL words was collected (appendix A). (Not included in the discussion here or in appendix A are words where variable secondary stress may attributed to preservation of stress from more than one word or a variable stress pattern in the embedded word – these are handled in §5.5.) Out of this total of 153 words, 15 words exhibited variable secondary stress. There were also two words where secondary stress is expected on the second syllable under preservation, but is consistently on the initial syllable: *Cù**nobelínus* (←*Cun**ó**belin*), and ***Tèr**psichoréan* (←*Terps**í**chore*). Altogether, 17/153, or 11.1%, of words from Wells exhibited evidence of relative prominence preservation failure.

The words with variable secondary stress collected from Wells are given in table 1. The pre-tonic syllable weight sequence is given, and it is indicated whether the same stress behaviour was exhibited in Jones (2003) for the word in question.

---

[9] In this statement, I admit sonorant-final syllables into the category of light syllable – just *envìronméntal* and *invòluntárily* from Wells (2000), and just *envìronméntal* from Jones (2003) – as the ability of sonorants to function as syllabic, therefore creating light syllables, is recognised, e.g. Pater (1995). The third syllables of these words are shown with reduced vowels in the pronouncing dictionaries, supporting this hypothesis (re: chapter 2). Neither *environmental* nor *involuntarily* are counted in the statistical calculations that test for relative prominence and second-syllable foot-head preservation (§4.4).

	Embedding word	Syllable weight sequence	Embedded word	Varies in Jones (2003)?
1	anfractuosity	#HHL	anfráctuous	---
2	exteriority	#HHL	extérior	---
3	somnambulation	#HHL	somnámbulate	---
4	Victoriana	#HHL	Victória	Y
5	antagonistic	#HLL	antágonist	---
6	anticipation	#HLL	antícipate	Y
7	anticipatory	#HLL	antícipate	---
8	Antipodean	#HLL	Antípodes	Y
9	concatenation	#HLL	concátenate	Y
10	(o)esophageal	#HLL	(o)esóphagus	Y
11	infinitival	#HLL	infínitive	---
12	Nicomachean	#HLL	Nicómachus	Y
13	participation	#HLL	partícipate	Y
14	prognostication	#HLL	prognósticate	N
15	Pythagorean	#HLL	Pythágoras	Y

Table 1: candidates for relative prominence preservation which have variable secondary stress in Wells (2000)

Similar behaviour was evident in the data from Jones (2003). A total of 174 words were collected, excluding words with consistently reduced vowels in their initial syllables (appendix B). Out of this total, 51 #HHL and #HHL words exhibited variable secondary stress; these are shown in appendix B. A further 8 words were found where second-syllable preservation was anticipated, but the dictionary showed secondary stress to be consistently upon the initial syllable, e.g. *rèpatriátion* (*repátriate*). These 8 words are shown in table 2 for illustration (as well as being included in appendix B).[10] Altogether, the proportion of words where relative

[10] Not included in this 8 are *dèactivátion*, *dècelerátion* and *rèallocátion*, as there is no corresponding base which has an initial heavy syllable (i.e. *deáctivate*, *decélerate*, *reállocate* are only given in the

prominence preservation fails some or all of the time in Jones (2003) stands at 59/174, or 33.9%.

	Embedding word	Syllable weight sequence	Embedded word
1	Bàrtolomméo	#HLL	Barthólomew
2	Tèrpsicoréan	#HLL	Terpsíchore
3	dèconsecrátion	#HHL	dècónsecrate
4	dèregulátion	#HLL	dèrégulate
5	prèdestinátion	#HLL	prèdéstine
6	rèactivátion	#HHL	rèáctivate
7	règenerátion	#HLL	règénerate
8	rèpatriátion	#HLL, #HHL	repátriate

Table 2: candidates for relative prominence preservation which have consistent initial-syllable secondary stress in Jones (2003)

4.2.2 Discussion

In chapter 3, we saw that two arguments had been made with respect to relative prominence preservation: (i) relative prominence preservation occurs without exception (Kiparsky, 1979); (ii) relative prominence preservation does not occur at all (Halle & Vergnaud, 1987a; Sainz, 1992). The data from §4.2.1 does not, at least on the face of things, support either hypothesis: while there are clear instances of relative prominence preservation failure, the majority of words from both Jones (2003) and Wells (2000) have secondary stress consistently on their second syllables, as expected under successful relative prominence preservation. Impressionistically, this situation resembles a new hypothesis that was proposed in chapter 3: that relative

dictionary with the short vowel [ɪ] in their initial syllables). It should also be noted that, aside from *Bàrtolomméo*, *Tèrpsichoréan* and *rèpatriátion*, all of the words in table 2 have bases where the initial prefixes are shown as bearing secondary stress in Jones (2003). This is not the case for many of the words on appendices A and B. It may be the case that productive prefixation is interfering with preservation here, as will be suggested in chapter 6.

prominence preservation does occur, but with variable success. This hypothesis is explored at length in chapter 6, and evidence is presented there which supports this proposal for variably successful preservation.

Interestingly, some of the data collected for §4.2.1 does not behave as we might expect with respect to the stressing of word-initial syllables. The argument for relative prominence preservation absolutely requires that the heavy initial syllables of embedding and embedded words are stressed, as well as their second syllables. However, some of the prefixed examples from Jones and Wells display reduced vowels in their initial syllables, particularly those examples with obscured Latinate prefixes: e.g. *contemporaneous*, *concelebration*, *configuration* and *conglomeration*. The tendency of Latinate prefixes, e.g. *ad-*, *ab-*, *con-*, *pre-*, *pro-*, *re-*, to occur with reduced vowels is well-recognised (e.g. SPE: 118; Liberman & Prince, 1977: 284). In §6.5.4.2, I discuss how the stresslessness of certain prefixes may interact with relative prominence preservation.

For now, in spite of the impression from our data that relative prominence preservation does occur to some extent, it remains to provide more solid evidence for the existence of relative prominence preservation. In §4.4, the secondary stress patterns of words which are candidates for relative prominence preservation are compared to those of phonotactically similar monomorphemic and bound-root base words. This comparison supports the very existence of relative prominence preservation, and also that of second-syllable foot-head preservation in heavy-initial words.

4.3 Monomorphemic and bound-root base words

The data from §4.2 does not give the impression that relative prominence preservation never occurs at all. Nevertheless, to have real evidence for relative prominence preservation, the behaviour of complex words must be compared to that of a control group of words which are not candidates for preservation, and any differences in behaviour evaluated. The data for words which are not candidates for preservation is presented here, and their secondary stress behaviour is evaluated against the predictions for monomorphemic and bound-root base words outlined in chapter 3.

To keep the data comparable to that collected for relative prominence preservation in §4.2, only monomorphemic words with pre-tonic #HLL and #HHL sequences were collected (if the third syllable was variably heavy, e.g. *praseodymium* which sometimes has a diphthong in its third syllable, the example was included in the data set to prevent it becoming even smaller). The monomorphemic and bound-root base words were subject to the same restrictions as complex words: if the stress pattern of the monomorphemic word varied but the same segmental representation did not occur with both stress variants, the example was excluded. It has also been indicated if words have a consistently reduced vowel in a syllable.

For the data from Wells (2000), 3 out of a total of 18 monomorphemic or bound-root base #HHL and #HLL words exhibited variable secondary stress placement, and two had consistent second-syllable stress. The full list of 18 words is given in table 3. To indicate a syllable with a consistently reduced vowel, '=R' follows the relevant syllable in brackets; the use of L/H for the third syllable indicates that the syllable sometimes has a long vowel which makes it heavy.

	Monomorphemic or bound-root base word	Syllable weight sequence	Secondary stress behaviour
1	chionodoxa	#HLL	Variable
2	coccidiosis	#HLL	Variable
3	Ticonderoga	#HHL	Variable
4	Àntofagásta	#H.L(=R).L	initial only
5	Bàndaranáike	#H.L(=R).L	initial only
6	càlceolária	#HLL	initial only
7	Càlvocoréssi	#H.L(=R).L	initial only
8	Còpacabána	#H.L(=R).L	initial only
9	Gùadalajára	#H.L(=R).L	initial only
10	Hèliogábalus	#H.L(=R).L/H	initial only
11	ònkaparínga	#H.L(=R).L	initial only
12	Pàpiaménto	#HLL	initial only

	Monomorphemic or bound-root base word	Syllable weight sequence	Secondary stress behaviour
13	pròprioceptor	#H.L.L/H	initial only
14	pùsillanímity	#HLL	initial only
15	tsùtsugamúshi	#H.L(=R).L	initial only
16	Yòknapatáwpha	#H.L(=R).L	initial only
17	embàrcadéro	#HHL	second only
18	Risòrgiménto	#HHL	second only

Table 3: behaviour of monomorphemic and bound-root base words from Wells (2000)[11]

The behaviour of the monomorphemic and bound-root base words in table 3 corresponds well with what we predicted on the grounds of the literature reviewed in chapter 3. The majority of #HLL words have consistent initial-syllable secondary stress – 13/15 words. Just two have variable secondary stress. Of the 15 #HLL words, several also have consistently reduced vowels in their second syllables. Taken together, these observations fit in with Hammond's (1989) hypothesis that #HLL words have the default footing #(HL)L (§3.1).

It is harder to deduce a picture for #HHL words – there are just 3 examples. The well-worn example *Ticonderoga* has variable secondary stress; however, both *emb**à**rcadéro* and *Ris**òr**giménto* have consistent second-syllable secondary stress. The lack of variation in these examples is not necessarily problematic – Halle & Vergnaud (1987a: 245) suggested that there might be accidental gaps or "oversights" in words which are candidates for variable relative prominence over their first two syllables (§3.2.1). This may therefore account for why not all #HHL monomorphemic words having both initial- and second-syllable stressed variants.

The picture from Jones (2003) for monomorphemic and bound-root base words is similar. The examples are given in table 4.

[11] ***Ar**boriculture* and ***cam**pylobacter* have both been removed. These words have alternative stress patterns which suggest compound behaviour: *árboricùlture* and *cámpylobàcter*.

	Monomorphemic or bound-root base word	Syllable weight sequence	Stress behaviour
1	icosahedron	#HLL	variable
2	Àntofagásta	#H.L(=R).L	initial only
3	Bàndaranáike	#H.L(=R).L	initial only
4	Bàrquisiméto	#HLL	initial only
5	càlceolária	#HLL	initial only
6	Cònstantinóple	#HHL	initial only
7	Còpacabána	#H.L(=R).L	initial only
8	Hèliogábalus	#H.L.L/H	initial only
9	Pènthesiléa	#HLL	initial only
10	pràseodýmium	#H.L.L/H	initial only
11	pròprioception	#H.L.L/H	initial only
12	pùsillanímity	#HLL	initial only
13	Ahàsuérus	#HLL	second only

Table 4: behaviour of monomorphemic and bound-root base words from Jones (2003)

We cannot really make any generalisations for the behaviour of #HHL words from Jones (2003): there is just one example, ***Còn**stantinóple*. For the other 12 words, all of which have #HLL, the picture is again as we would expect: 10/12 words have consistent initial-syllable secondary stress, with just one having variable secondary stress, and one consistent second-syllable secondary stress.

4.4 Support for second-syllable preservation

We can now compare the behaviour of words where second syllable relative prominence preservation is anticipated (§4.2) to that of those which are not candidates for stress preservation (§4.3), and establish the strength of the evidence for both relative prominence preservation and second-syllable foot-head preservation. This examination will only consider #HLL words both categories: there are so few #HHL monomorphemic and bound-root base words from both Jones and Wells that there are not enough for any meaningful comparison.

As we noted in §3.4, we can examine #HLL words where second-syllable preservation is anticipated not only to determine evidence for relative prominence preservation, but also for evidence of second-syllable foot-head preservation. This is because, as argued by Hammond (1989), and as supported by the evidence here, the second syllables of #HLL words are not usually the head of a foot in the absence of second-syllable stress preservation. Consequently, variable secondary stress cannot occur in such words. In contrast, even if second-syllable relative prominence preservation fails, second-syllable foot-head preservation would result in the second syllable of the complex word being the head of a foot, and variable stress secondary stress would therefore have the opportunity to occur. Evidence for second-syllable foot-head preservation may therefore come from a relative absence of variable, as well as consistent second-syllable, secondary stress in monomorphemic and bound-root base words as compared to those where second-syllable preservation is anticipated.[12]

To try and ensure that all of the words analysed here do indeed have light second and third syllables, and are therefore incontrovertibly #HLL, some particularly controversial cases of syllable weight have been factored out: words with /sC/ clusters following their second or third syllable; those with orthographic geminates, also in the same locations; words where the second or third syllables are sonorant-final; and, finally, words where the dictionary indicates the weight of the second and/or third syllable to vary between heavy and light. Two of these factors, /sC/ clusters and orthographic geminates, require further discussion.

In the initial classification of the data, it has been assumed that /sC/ clusters syllabify as the onset of the second syllable, following Selkirk (1982b: 360), and also Kahn (1976). This proposal is by no means uncontroversial. Kager (1989: 117-118) argues that whether /s/ is treated as the coda of a preceding syllable or the onset of a

[12] Admittedly, this may only be an approximate indication of the existence of second-syllable foot-head preservation: it cannot be assumed that the placement of secondary stress will vary in every word where it could, particularly in light of Halle & Vergnaud's (1987a) argument for accidental gaps (§3.2.1).

following syllable in a cluster context “is lexically marked to a certain extent”. This can be demonstrated by the examples in (2) and (3). In (2), /s/ is treated as closing the words’ second syllables, thus causing them to receive weight-sensitive primary stress:

(2) /s/ functions as a coda (Kager, 1989: 118)

orch**és**tral	phlog**ís**ton	ill**ús**trative
ab**ýs**mal	ill**ús**trate	conf**ís**catory

In contrast, for the examples in (3), Kager argues that /s/ syllabifies as an onset to their third syllables. Consequently, their second syllables are light, and are skipped over for primary stress:

(3) /s/ functions as an onset (Kager, 1989: 117)

órch**es**tra	mág**is**tral	mín**is**trant	prót**es**tant
péd**es**tal	mín**is**ter	rég**is**trant	tál**is**man

Burzio (1994: 61) and Sainz (1992: 60-71) take a different approach to /sC/ clusters, but to the same end. Both argue that the /s/ of an /sC/ cluster is always syllabified in the coda of the preceding syllable, but that syllables closed by /s/ are idiosyncratically light or heavy. Given the observed unpredictable effect of /sC/ clusters upon syllable weight, words containing them in pertinent positions have been excluded from the following chi-square analysis – these words cannot be said to be incontrovertibly #HLL.

Geminate clusters are similarly contentious in terms of syllable weight: although English is generally argued not to have any form of distinctive consonant length, there is evidence to suggest that syllables associated with orthographic geminates behave as heavy, not light, as indicated by the exceptional penultimate stress in *rub**éll**a*, *Mad**ónn**a* and *conf**étt**i* (SPE: 148; Rice, 1996; Giegerich, 1999: 164). Therefore, any words have been excluded from the following chi-square analysis where the weight of one of the word’s first three syllables is determined by the effects of orthographic geminates, as, again, these words cannot be said to be incontrovertibly #HLL.

Words with consistent schwas in their first or second syllables were not excluded from the analysis. Although, under Sainz's (1992) proposal (chapter 3), a consistent schwa could be argued to interfere with stress assignment, this decision was made so as to control complex and monomorphemic words to a similar extent: it was necessary to include #H.L(=R).L monomorphemic and bound-root base words in order to have enough data for a potentially reliable chi-square analysis.

For Wells (2000), a total of 70 #HLL words which were candidates for relative prominence preservation were examined; of these, 10 had variable secondary stress, 1 had consistent initial-syllable secondary stress, and 59 had consistent second-syllable secondary stress. A total of 13 #HLL monomorphemic or bound-root base words from Wells (2000) were examined – 2 words were removed where the weight of the second or third syllables could be argued to be heavy. Of the 13 words, two had variable secondary stress, and the rest had consistent initial-syllable secondary stress.

For Wells (2000), there is highly significant statistical support in a chi-square analysis for both relative prominence preservation and second-syllable foot-head preservation: $\chi^2 = 37.893$, 1df, $p<.001$ for the former; $\chi^2 = 61.346$, 1df, $p<.001$ for the latter.[13] However, it should be noted that both results must be treated with some caution, as there is an expected frequency less than 5 in each analysis.[14] Fisher's exact test and the likelihood ratio can both be used in such an event, and these are also both very significant with respect to both foot-head and relative prominence preservation. The chi-square contingency tables, along with all of the statistical results, are given in (4).

[13] Chi-square is a statistical test that can tell us the likelihood of a relationship between categorical variables: in this case, whether the word is a candidate for stress preservation or not, and the word's secondary stress behaviour. All chi-square results given here are without Yates' correction. Yates' correction is an adjustment made to the chi-square test in order to ensure that statistical significance is not overestimated in analyses of small amounts of data; however, the test often overcorrects (Field, 2005: 686).

[14] Expected frequencies below 5 mean the chi-square result is unreliable.

(4) Second-syllable preservation with #HLL in Wells (2000)

(a) Relative prominence preservation

Wells (2000)	Candidate for relative prominence preservation?	
	Complex, expected on second syllable	Monomorphemic / not a candidate
Variable or consistent initial-syllable stress with #HLL	11 (20.2 expected) 15.7%	13 (3.8 expected) 100%
Consistent second syllable stress with #HLL	59 (49.8 expected) 84.3%	0 (9.2 expected) 0%

χ^2 = 37.893, 1df, p<.001; 1 cell has an expected frequency less than 5[15]

Fisher's exact test: p<.001, 2-tailed; Likelihood ratio = 38.945, 1df, p<.001

[15] An 'expected frequency' is the frequency we might get by chance, if there was no relationship between the variables. In a contingency table of this size, the chi-square test is only reliable if all expected frequencies are greater than 5.

(b) Second-syllable foot-head preservation

Wells (2000)	Candidate for foot-head preservation?	
	Complex, expected on second syllable	Monomorphemic / not a candidate
Variable or consistent second-syllable secondary stress with #HLL	69 (59.9 expected) 98.6%	2 (11.1 expected) 15.4%
Consistent initial-syllable stress with #HLL	1 (10.1 expected) 1.4%	11 (1.9 expected) 84.6%

χ^2 = 61.346, 1df, p<.001; 1 cell has an expected frequency less than 5
Fisher's exact test: p<.001, 2-tailed; Likelihood ratio = 46.944, 1df, p<.001

In the analysis of the data from Jones (2003), all of the same syllable-weight controls were applied, except that words where the weight of the third syllable varies between heavy and light were not excluded for either the complex category or the monomorphemic and bound-base category. This was because, if such examples were excluded from the monomorphemic and bound-root base set (*prò̀priocéption*, *prà̀seodýmium* and *Hè̀liogábalus*), there would be just 8 words from this morphological group for analysis.[16] There were therefore 11 monomorphemic and bound-root base words: 1 with consistent second-syllable secondary stress; 1 with variable secondary stress; and 9 with consistent initial-syllable secondary stress. For the complex words, there were 80 words in all. Of these, 27 had variable secondary stress, 4 had had consistent initial-syllable secondary stress, and 49 had consistent second-syllable secondary stress.

[16] Fisher's exact test may be used with less than 10 cases in a category, but a minimum of 10 is required for a chi-square test. As I always use chi-square in the first instances here (the total sample size in each 2x2 contingency table is in all cases greater than 20), I stick to this minimum of 10 cases.

For Jones (2003), the picture was similar to that from Wells in that there was significant support for both relative prominence and second-syllable foot-head preservation. For relative prominence preservation, the result of the chi-square analysis was $\chi^2 = 10.627$, 1df, p = .001. For foot-head preservation, the result was $\chi^2 = 46.603$, 1df, p<.001. However, the result for foot-head preservation should be treated with some caution, as one of the expected frequencies is below 5. Happily, there are significant results with both Fisher's exact test and the likelihood ratio for foot-head preservation. The contingency tables for the chi-square analyses, along with all of the statistical results, are given in (5).

(5) Second-syllable preservation with #HLL in Jones (2003)

(a) Relative prominence preservation

Jones (2003)	Candidate for relative prominence preservation?	
	Complex, expected on second syllable	Monomorphemic / not a candidate
Variable or initial-syllable stress with #HLL	31 (36.0 expected) 38.8%	10 (5.0 expected) 90.9%
Consistent second syllable stress with #HLL	49 (44.0 expected) 61.3%	1 (6.0 expected) 9.1%

$\chi^2 = 10.627$, 1df, p = .001

(b) Second-syllable foot-head preservation

Jones (2003)	Candidate for foot-head preservation?	
	Complex, expected on second syllable	Monomorphemic / not a candidate
Variable or consistent second-syllable secondary stress with #HLL	76 (68.6 expected) 95.0%	2 (9.4 expected) 18.2%
Consistent initial-syllable stress with #HLL	4 (11.4 expected) 5.0%	9 (1.6 expected) 81.8%

χ^2 = 46.603, 1df, p<.001. 1 cell has an expected count less than 5

Fisher's exact test: p<.001, 2-tailed; Likelihood ratio = 32.448, 1df, p<.001

In conclusion, we have significant statistical support for both relative prominence preservation and second-syllable foot-head preservation from both Jones (2003) and Wells (2000). Relative prominence preservation is not consistently successful, contrary to Kiparsky's (1979) proposal, but there is nevertheless strong evidence for its existence. And, even when relative prominence preservation fails, complex words do not always behave like monomorphemic or bound-root base forms: second-syllable foot-head preservation may still occur.

4.5 Initial-syllable foot-head preservation

We noted in chapter 3 that we do not expect heavy-initial words where initial-syllable preservation is anticipated, e.g. *gentrification* (*géntrify*), to be candidates for relative prominence preservation. This is because, in the relevant embedded words, only the initial syllable and not the second is expected to be stressed; there is therefore no relative prominence between stresses to preserve. Data for heavy-initial words which are candidates for initial-syllable preservation was collected to test for evidence for initial-syllable foot-head preservation.

A total of 127 heavy-initial words were collected from Wells (2000), of which one – *phantasmagoria* – had variable secondary stress (all words from Wells are given in appendix C). That any word has variable secondary stress at all would appear to be instance of failure of foot-head preservation. However, there is an alternative explanation for *phantasmagoria* that indicates that this may be an exceptional case of a failure to preserve relative prominence in a word where initial-syllable preservation is expected. The second syllable is footed in *phantasmagoria*'s embedded word, *phánt**à**sm* (Wells, 2000). It may therefore be argued that *(**phan)(ta.**sma)(go.ri)<a>* preserves the foot structure of *(**phan)(ta**.sm),* and so is a candidate for variable placement of secondary stress under the failure of relative prominence preservation. The pronunciation *phantàsmagória* would therefore be an instance of failure to preserve relative prominence

For Jones (2003), the picture was similar. Out of a total of 116 words, two had variable secondary stress: *comparability* and *phantasmagoria* (all words from Jones are given in appendix D). *Phantasmagoria* may be subject to the same explanation as before, although, admittedly, the pronunciation *phántàsm*, with second-syllable secondary stress, is not given in Jones. *Comparability*, however, is an interesting case: it may not be an instance of preservation failure after all. Jones (2003) only gives *cómparable* for British English, but, for myself at least, there is a semantic difference between *compárable*, meaning 'can be compared', and *cómparable*, meaning 'equal'. (Both pronunciations of *comparable* are given in Wells (2000) for British English.) Burzio (2005b) appears to note the same phenomenon:

> *[C]ómparable*, which is different in stress from its base *compáre*, is also divergent from it semantically, while *compárable*, which retains the stress of the verb, is also strictly faithful to the verb semantically (Burzio, 2005b: 68).

And, similarly to Burzio, Raffelsiefen (1993: 34-5) suggests that *compáre* is not the "synchronic base of *cómparable*", but that *cómparable* is "synchronically baseless"; in contrast, *compárable* may be argued to have the synchronic base *compáre*. With two different embedded words, there is no 'variation in *comparability*': ***còm**parabílity* (*(**com**.pa)ra(bí.li)<ty>* ← *(**cóm**.pa)<rable>*) and *comp**à**rabílity* (*(com)(**pa**.ra)(bí.li)<ty>* ← *(com)(**pá**.ra)<ble>* ← *(còm)(**páre**)*) are two different words by virtue of their different semantics and word-formation histories (see also Marvin, 2002: 75).[17]

4.5.1 Evidence for initial-syllable preservation

So far, the data impressionistically supports the hypothesis that there is initial-syllable preservation of foot heads in heavy-initial words: very few heavy-initial complex words which are candidates for initial-syllable preservation ever have secondary stress anywhere other than their initial syllables. It will now be shown that there is also statistical support for initial-syllable foot-head preservation.

To find the evidence for initial-syllable preservation, I compared the #HLL words from Wells and Jones with the control group of monomorphemic and bound-root base #HLL words from §4.3. #HHL words were not examined: there are very few such monomorphemic and bound-root base examples, and, as anticipated (§3.1, §3.4), there were also very few examples of #HHL words which were candidates for initial-syllable preservation. Consistently #HLH words were also not examined from either Jones (2003) or Wells (2000): the weight of the third syllable is variable in the great majority of #HLH words where initial-syllable preservation is anticipated (attributable to *-ize* formations, e.g. *stabilization* (*stábilize*), where the third syllable varies between [aɪ]~[ɪ]), leaving very few incontrovertibly #HLH words to examine.

As with second-syllable preservation, the data examined with respect to initial-syllable preservation has been controlled as far as possible. For Wells, all words were excluded where the weight of the second or third syllable varied. For Jones, as before, I was unable to exclude words where the weight of the third syllable varies between heavy and light: to do so would exclude too many monomorphemic and bound-root base words. The other steps taken for second-syllable preservation – removal of words /sC/ clusters, orthographic geminates, and sonorant-final syllables in pertinent positions – applied here too. As with second-syllable preservation, #HL(=R)L words were not excluded: to do so would exclude too many

[17] Ricardo Bermúdez-Otero (personal communication) has suggested that a variable relationship with the embedded word *compare* may not be the only cause of the variable primary stress in *comparable*, as there is variable primary stress in *formidable*, which is formed upon a bound root. Variable primary stress could, therefore, be a purely phonological phenomenon associated with *-able* suffixation.

monomorphemic and bound-root base examples. For Jones, *comparability* was excluded in light of the argument for variation in the embedded word.

There was statistical support for initial-syllable foot-head preservation from both Jones (2003) and Wells (2000). Chi-square tests were significant but unreliable in both cases; however, both Fisher's exact test and the likelihood ratio were significant for both Wells (2000) and Jones (2003). All of the statistical results are given in (6) and (7):

(6) Initial-syllable foot-head preservation with #HLL in Wells (2000)

Wells (2000)	Candidate for initial-syllable preservation?	
	Complex, expected on initial syllable	Monomorphemic / not a candidate
Variable or consistent second syllable secondary stress	0 (1.6 expected) 0.0%	2 (0.4 expected) 15.4%
Consistent initial-syllable stress	46 (44.4 expected) 100.0%	11 (12.6 expected) 84.6%

χ^2 = 7.325, 1df, p<.01; 2 cells have expected frequencies below 5

Fisher's exact test: p<.05 (2-tailed)

Likelihood ratio = 6.307, 1df, p<.05

(7) Initial-syllable foot-head preservation with #HLL in Jones (2003)

Jones (2003)	Candidate for initial-syllable preservation?	
	Complex, expected on initial syllable	Monomorphemic / not a candidate
Variable or consistent second syllable secondary stress	0 (1.8 expected) 0.0%	2 (0.2 expected) 18.2%
Consistent initial-syllable stress	111 (109.2 expected) 100.0%	9 (10.8 expected) 81.8%

$\chi^2 = 20.518$, 1df, $p<.001$; 2 cells have expected frequencies less than 5

Fisher's exact test: $p<.01$ (2-tailed)

Likelihood ratio = 9.979, 1df, $p<.005$

In conclusion, we have significant statistical support for initial-syllable foot-head preservation in heavy-initial words. Heavy-initial words which are candidates for initial-syllable preservation are not expected to be candidates for relative prominence preservation.

4.6 Conclusion

Two key empirical findings have been reported in this chapter:

(i) Evidence for stress preservation in words with heavy initial syllables.

(ii) An indication that relative prominence preservation is a variably successful phenomenon.

Statistical evidence was presented both for the existence of relative prominence preservation, and for initial- and second-syllable foot-head preservation. This finding is very important indeed: it means that a theory of English phonology **must** be able to handle English weak stress preservation. The necessity of a weak stress preservation mechanism will be reinforced in the next chapter by the evidence presented for the existence of left-edge foot-head preservation in light-initial words.

The indication that relative prominence preservation is a variably successful phenomenon is very interesting indeed. The next step with respect to relative prominence preservation is to discover what factors condition the variable success of relative prominence preservation. This investigation is carried out in chapter 6; there, it is shown that relative prominence preservation is a probabilistic phenomenon predicted by word frequency. In chapter 8, this frequency result will be shown to have important implications for the precise type of weak stress preservation mechanism that is required.

Chapter five: the data for left-edge stress preservation II: light-initial words and other types of left-edge preservation

5.0 Introduction

In chapter four, it was shown that there is evidence for both relative prominence and foot-head preservation in words beginning with heavy initial syllables; results were also given which indicated that relative prominence preservation may be a variably successful phenomenon. In this chapter, similar results are presented for foot-head preservation in words with light initial syllables. It is shown that there is evidence for left-edge foot-head preservation, but that, as in chapter 4, preservation does not appear to be successful one-hundred percent of the time (§5.1-§5.4). Additionally, some more marginal instances of preservation are considered (§5.5): embedding words which may preserve stress from more than one embedded word or preserve the variable stress pattern of a single embedded word, and the preservation of metrical structure other than primary stress. The chapter is concluded in §5.6.

The methods of data collection and categorisation that were outlined in §4.1 are also applicable to the data presented in this chapter.

5.1 Words where second-syllable preservation is expected

The data for second-syllable foot-head preservation in light-initial words is presented in this section.

In chapter 3, it was shown that words with light initial syllables are expected to be candidates only for left-edge foot-head preservation, not relative prominence preservation. This is because, in embedding words like *originality* (*o**rí**ginal*) and *animalistic* (***á**nimal*), only the first **or** second syllable will ever be the head of a foot under preservation: *o(rí.gi)<nal>* → *o(rì.gi)nálity*; *(á.ni)<mal>* → *(à.ni)malístic*. Consequently, there is no opportunity for variable relative prominence – secondary stress – between the first and second syllables of the embedding word, because only one of the syllables will be the head of a foot in the first place.

5.1.1 The data

The data from Jones (2003) and Wells (2000) consists of words where the second syllable is shown to bear primary stress in the embedded word and the word-initial syllable is light, e.g. *originality* (*orı́ginal*).

Following from the hypotheses laid out in §3.4.1, both #LLL and #LLH words are expected to provide evidence for second-syllable preservation. However, in practice, #LLH words are very infrequent in the data collected from the pronouncing dictionaries: just *devèlopméntal* (*devélop*) and variably *machìcolátion* (*machícolate*) from both Wells (2000) and Jones (2003). #LLH words are henceforth ignored in the analysis on the grounds of insufficient data.

The other two possible trisyllabic pre-tonic sequences that begin with a light syllable - #LHL and #LHH – are not anticipated to provide evidence for second-syllable preservation (§3.4.1). Weight-sensitive stress assignment is independently expected to assign stress to the second syllables of #LHL words, as evident in monomorphemic *Mon**òn**gahéla*. If this is indeed the case, then it will not be possible to distinguish between behaviour caused by preservation, and that which results from purely phonological stress assignment in this phonological environment. #LHH words are not anticipated to provide evidence for preservation for a different reason: following the literature review presented in chapter 3, we do not anticipate finding many, if any, examples of #LHH monomorphemic and bound-root base words with which to compare the behaviour of words which are candidates for preservation. #LHL and #LHH words which are also candidates for second-syllable preservation are briefly examined here, but only to check for any behaviour that contradicts the hypotheses outlined in chapter 3.

There are #LLL words where second-syllable foot-head preservation sometimes fails; these are given in table 1 for both Jones (2003) and Wells (2000). Where a word is not in the dictionary in question, this is indicated by a line in the relevant box.

	Embedding word	Embedded word	Secondary stress of embedding word	
			Jones (2003)	Wells (2000)
1	anastomosis	anástomose	initial	variable
2	Boliviano	Bolívia	variable	second
3	discolouration	discólour	variable	___
4	dissatisfaction	dissátisfy	___	variable
5	dissimilarity	dissímilar	second	variable
6	dissimilation	dissímilate	second	variable
7	Hispaniola	Hispánic	variable	variable
8	horripilation	horrípilate	variable	variable
9	illegibility	illégible	variable	second
10	inamorato	inámorate	variable	variable
11	irregularity	irrégular	second	variable
12	miscegenation	miscégenate	initial	variable
13	vaticination	vatícinate	variable	variable

Table 1: #LLL words where second-syllable foot-head preservation fails

For Jones (2003), there were a total of 88 #LLL words where second-syllable foot-head preservation is expected, not including words with a consistently reduced vowel in their initial syllable. Out of this total, 9 words, or 10.2%, failed to exhibit consistent second-syllable foot-head preservation. The proportion of failure for Wells (2000) was similar: 10 out of a total of 87 #LLL words, or 11.5%, failed to exhibit consistent second-syllable foot-head preservation. The full lists of #LLL (and #LLH) words which are candidates for second-syllable preservation are given in appendices E and F for Wells (2000) and Jones (2003) respectively.

#LHL and #LHH words from both Wells (2000) and Jones (2003) are given on appendices G and H; just one word from each dictionary, *aràchnophóbia*, ever occurs with the syllable weight sequence #LHH. Interestingly, a very small number of #LHL words appear to defy the predictions of both phonology (cf.

monomorphemic *Mon**ò**n**gahéla*) and preservation: they sometimes occur with initial-syllable secondary stress. In Wells (2000), *inauguration* (*in**áu**gurate*), *posteriority* (*post**é**rior*) and *listeriosis* (*list**é**ria*) occur with variable secondary stress; in Jones (2003), *listeriosis*. Excluding words with consistent schwa in their initial syllable, 3/61 or 4.9% of #LHL words from Wells exhibit failure of second-syllable foot-head preservation, also defying purely phonological predictions; the figure is 1/54 or 1.9% for Jones.

5.1.2 Discussion

The data for second-syllable foot-head preservation in #LLL words indicates that this type of preservation is not completely successful all of the time, but is a variably successful phenomenon. This matches the observation made for relative prominence preservation in chapter 4 – relative prominence preservation was also not uniformly successful.

The proportion of failure of second-syllable foot-head preservation is very marginally greater than that of relative prominence preservation for Wells (2000): 11.5%, as compared to 11.1%. For Jones (2003), there is a marked difference in the opposite direction: the proportion of failure is 10.2% for second-syllable foot-head preservation, versus 33.9% for relative prominence preservation. The difference for Jones is significant in a chi-square analysis (χ^2= 17.054, 1df, p<.001), as shown in (1).

(1) Relative prominence versus second-syllable foot-head preservation in Jones (2003)

Jones (2003)		Type of second-syllable preservation	
		Relative prominence (#HLL and #HHL words)	Foot head (#LLL words)
Secondary stress	Consistent second-syllable stress	115 (128.8 expected) 66.1%	79 (65.2 expected) 89.8%
	Variable or consistent initial-syllable stress	59 (45.2 expected) 33.9%	9 (22.8 expected) 10.2%

χ^2= 17.054, 1df, p<.001

Possible causes for this apparent greater fallibility of relative prominence preservation as compared to second-syllable foot-head preservation are considered in §6.5.1 and §9.2.2.

With respect to the three #LHL words with variable secondary stress – *posteriority*, *listeriosis* and *inauguration* – it is of interest that, out of these three exceptions, two have an /sC/ cluster following their first syllable: it was noted in §4.4 that /sC/ clusters have been argued to cause the syllables they follow to be heavy. If the pre-tonic syllable weight sequence of *posteriority* and *listeriosis* is #HHL, not #LHL, as a result of the /sC/ cluster making the initial syllable heavy, then variable secondary stress is not unexpected from a purely phonological perspective – variable secondary stress is seen in the monomorphemic #HHL word *Tìconderóga~Ticònderóga*. When these words have initial-syllable secondary stress words, it could therefore be an instance of failure to preserve relative prominence rather than failure of even foot-head preservation. An explanation for *inauguration* is less obvious, but the fact that this word contains a prefix is of interest. Giegerich (1999: 239-40) argues that in the lexical phonology (as opposed to the postlexical phonology), syllabification never takes place across prefix boundaries except in the case of cranberry roots (e.g. cranberry-based *inépt* would lexically syllabify as [ɪ.nɛpt], not *[ɪn.ɛpt]).[1] Wells (2000) gives *augur*, which may be argued to be the free root [-*augur*-]$_R$ in *inauguration* (both *augur*$_V$ and *inaugurate*$_V$ can mean 'to induct into office' according to the *OED Online*). We could therefore argue that the [n] of the prefix in *inauguration* syllabifies as a coda, rather than with the following [ɔː]; consequently, the pre-tonic syllable weight sequence of *inauguration* would be #HHL, not #LHL. Again, in light of variable secondary stress in #HHL *Tìconderóga~Ticònderóga*, initial-syllable stress would no longer defy purely phonological predictions.

5.2 Monomorphemic and bound-root base words

In this section, the secondary stress behaviour of light-initial monomorphemic and bound-base words with three pre-tonic syllables is explored. Any contrast between the behaviour of these words and those which are candidates for second-syllable foot-head preservation (§5.1) will provide evidence for the existence for preservation; this evidence is presented in §5.3. The monomorphemic and bound-root base words presented here will also be used to provide evidence for initial-syllable preservation in §5.4.

As was discussed in §3.4.1 and §5.1, only #LLL and #LLH monomorphemic and bound-root base words are anticipated to be useful in finding evidence for second-syllable foot-head preservation: phonological predictions overlap with those of second-syllable preservation with #LHL, and we do not expect to find many, if any, examples of #LHH words. Furthermore, #LHL monomorphemic and bound-root base words are not expected be to be useful in finding evidence for initial-syllable preservation: it was argued in §3.4.1 that, due to the weight-sensitivity of stress assignment to the embedded word, candidates for initial-syllable preservation are unlikely to occur with the syllable-weight sequence #LHL. Nevertheless, data for any #LHL and #LHH monomorphemic and bound-root base words is presented here to verify that it meets the predictions made in chapter 3.

---

[1] Hurrell (2001) offers a detailed exploration of stratum one prefix bracketing. Examples of prefixes syllabifying separately with bound and free roots are given by Hurrell (2001: 72).

5.2.1 The data

The data for #LLL and #LLH words are presented in table 2 for both Wells (2000) and Jones (2003):

	Word	Secondary stress		Pre-tonic syllable weight sequence	
		Wells (2000)	Jones (2003)	Wells (2000)	Jones (2003)
1	Arimathaea	initial	initial	#LLL	#LLL
2	Ballymacarrett	initial	---	#LLL	---
3	Cassiopeia	initial	initial	#LLH	#L.L.H/L
4	Cassivelaunus	initial	initial	#LLL	#LLL
5	Coriolanus	initial	initial	#LLL	#L.L.H/L
6	Czechoslovakia[2]	initial	initial	#L.L.H/L	#L.L.H/L
7	dolichosaurus	initial	---	#LLH	---
8	enneahedron	initial	---	#LLL	---
9	Halicarnassus	initial	initial	#LLH	#LLH
10	Kilimanjaro	initial	initial	#LLH	#LLH
11	Liliburlero	initial	initial	#LLL	#LLL
12	liriodendron	initial	---	#L.L.H/L	---
13	Machiavelli	initial	initial	#LLL	#LLL
14	Mediterranean	initial	initial	#LLL	#LLL
15	Mulligatawny	initial	initial	#LLL	#LLL
16	Nebuchadnezzar	initial	initial	#LLH	#L.L(=R).H
17	Papiamento	initial	---	#LLL	---
18	peripatetic	initial	initial	#LLL	#LLL
19	peripeteia	initial	---	#LLL	---
20	peritoneum	initial	initial	#L.L.H/L	#LLL
21	pichiciego	initial	---	#LLL	---

[2] *Czechoslovakia* is not umambiguously monomorphemic or bound-root based – cf. *Czech*. I have chosen to include *Czechoslovakia* here given the extreme differences in length of this word and *Czech* – the two words share relatively little segmental material.

	Word	Secondary stress		Pre-tonic syllable weight sequence	
		Wells (2000)	Jones (2003)	Wells (2000)	Jones (2003)
22	plenipotentiary	initial	initial	#L.L.H/L	#LLL
23	Torremolinos	initial	initial	#LLL	#LLL
24	trichomoniasis	initial	---	#L.L.H/L	---
25	Winnepesaukee	initial	---	#LLL	---
26	Abergavenny	initial	initial	#L.L(=R).L	#L.L(=R).L
27	Abertillery	initial	initial	#L.L(=R).L	#L.L(=R).L
28	abracadabra	initial	initial	#L.L(=R).L	#L.L(=R).L
29	Alamagordo	initial	initial	#L.L(=R).L	#L.L(=R).L
30	anacoluthon	initial	initial	#L.L(=R).L	#L.L(=R).L
31	Asarabacca	initial	---	#L.L(=R).L	---
32	coloratura (≠ 'colour')	initial	initial	#L.L(=R).L	#L.L(=R).L
33	Estremadura	---	initial	---	#L.L(=R).L
34	flibbertigibbet	initial	initial	#L.L(=R).L	#L.L(=R).L
35	Guadalajara	---	initial	---	#L.L(=R).L
36	memorabilia (≠ 'memory')	initial	initial	#L.L(=R).L	#L.L(=R).L
37	paraphernalia	initial	initial	#L.L(=R).L	#L.L(=R).L
38	prosopopeia	initial	---	#L.L(=R).L/H	---
39	Savonarola	initial	initial	#L.L(=R).L	#L.L(=R).L
40	tatterdemalion	initial	initial	#L.L(=R).L	#L.L(=R).L
41	Bucaramanga	---	second	---	#LLL
42	Byelorussia	second	---	#LLL	---
43	egalitarian	second	second	#LLL	#LLL
44	Ekaterinburg	---	second	---	#LLL
45	Epaminondas	second	second	#LLL	#LLL
46	Scheherazade	second	second	#LLL	#LLL

	Word	Secondary stress		Pre-tonic syllable weight sequence	
		Wells (2000)	Jones (2003)	Wells (2000)	Jones (2003)
47	acciaccatura	second	second	#L(=R).L.L	#L(=R).L.L
48	Ahasuerus	second	second	#L(=R).L.L	#L(=R).L.L
49	Ahenobarbus	second	second	#L(=R).L.H/L	#L(=R).L.H/L
50	amanuensis	second	second	#L(=R).L.L	#L(=R).L.L
51	anachronistic	second	second	#L(=R).L.L	#L(=R).L.L
52	Apollinaris	second	second	#L(=R).L.L	#L(=R).L.L
53	apotheosis	variable	variable	#LLL	#LLL
54	episiotomy	variable	variable	#LLL	#LLL
55	Iphigeneia	variable	variable	#LLL	#LLL
56	Navratilova	variable	variable	#LLL	#LLL

Table 2: secondary stress behaviour of monomorphemic and bound-base #LLL and #LLH words

A majority of #LLL and #LLH words from both Wells (2000) and Jones (2003) exhibit consistent initial-syllable secondary stress, as in *àbracadábra* and *Hàlicarnássus*: excluding words with consistently reduced vowels in their first or second syllables, 25/33 words (75.8%) from Wells (2000), and 16/25 words (64.0%) from Jones (2003), have consistent initial-syllable secondary stress. Consistent second-syllable stress occurs in 4/33 words (12.1%) from Wells, and 5/25 words (20.0%) from Jones (2003). Variable secondary stress occurs in 4/33 (12.1%) words from Wells (2000), and 4/25 (16.0%) words from Jones (2003).

Examples of #LHL monomorphemic and bound-root base words were also found. In chapter three, it was predicted that monomorphemic or bound-root base words with #LHL will have consistent second-syllable secondary stress, as in *Mon**òn**gahéla*. This hypothesis is quite well-supported: 8 out of the 10 #LHL words found have consistent second-syllable secondary stress, as shown in table 3:

	Word	Secondary stress	
		Wells (2000)	Jones (2003)
1	aggiornamento	second	second
2	Ahenobarbus	second	---
3	Balenciaga	second	---
4	chiaroscuro	second	second
5	episiotomy	second	second
6	Monongahela	second	---
7	Scheherazade	second	second
8	Tegucigalpa	second	second
9	Amontillado	variable	second
10	apotheosis[3]	---	initial

Table 3: secondary stress behaviour of monomorphemic and bound-root base #LHL words

Nevertheless, there are two exceptions to the behaviour anticipated for #LHL monomorphemic and bound-base words: variable secondary stress in *Amontillado*, and consistent initial-syllable stress in *apotheosis*.

Finally, a small number of examples of #LHH monomorphemic and bound-root base words were found. These are given in table 4:

[3] With the pronunciation [æ.pəʊ]*theosis* given in Jones (2003).

Word	Stress	
	Wells (2000)	Jones (2003)
Bophuthatswana	Initial	initial
Czechoslovakia[4]	Initial	initial
prosopopoeia	Initial	---
trichomoniasis	Initial	---

Table 4: secondary stress behaviour of monomorphemic and bound-base #LHH words

No hypotheses were made for the behaviour of #LHH monomorphemic and bound-base words in chapter 3 due to insufficient discussion of this pre-tonic syllable-weight sequence in past literature. However, it is worth noting the uniform secondary stress behaviour among the words in table 4 – all words have consistent initial-syllable stress.

5.2.2 Discussion

The data for monomorphemic and bound-root base #LLL and #LLH words is in keeping with our hypotheses from chapter 3: the majority of words have consistent initial-syllable secondary stress, as in *àbracadábra*. However, the words which do not have this default pattern, but which have second-syllable secondary stress some or all of the time instead, are interesting: although they are not a clearly defined group of exceptions, they do have some phonological characteristics in common to which their second-syllable secondary stress may be attributable.

As discussed in §3.3.2, Halle & Kenstowicz (1991) proposed that the group of #LLL words with second- rather than initial-syllable secondary stress (e.g. *Epàminóndas*) have the defining characteristics of being (i) etymologically Greek, and (ii) having word-initial syllables that lack onsets. The etymological argument is certainly true for some of the examples which have second-syllable secondary stress some or all of the time: e.g. *Epàminóndas*, *Ahènobárbus*, and the second-stressed

[4] *Czechoslovakia* is classified as #LHH here because the dictionaries report a possible pronunciation with [əʊ] in the word's second syllable.

variant *Iphìgenéia*. However, there are also exceptions: e.g. *Bucàramánga*, and the second-stressed variant *Navràtilóva*. Examples like *Bucàramánga* and *Navràtilóva* also cast doubt on Halle & Kenstowicz's argument that second-syllable secondary stress will only occur in #LLL words with onset-less initial syllables (see also Pater, 1995: f.n. 11). Nevertheless, there does appear to be something in Halle & Kenstowicz's argument with regards to their observation about word-initial syllable onsets. Out of all #LLL words with consistent second-syllable secondary stress or a second-syllable stressed variant in table 2, 12/16 (75%) have a word-initial syllable which lacks an onset; this compares to just 9/35 (25.7%) of words with consistent initial-syllable secondary stress.[5] This is a statistically significant difference in a chi-square analysis (χ^2 = 11.012, 1df, p= .001), and is a tendency also noted by Coleman (2000: 177).

It is plausible that a syllable's lack of an onset may cause it to reject stress. Stress assignment is generally argued to be insensitive to onsets (onsets not counting towards syllable weight), but both Nanni (1977) and Davis (1988) argue for onset-sensitivity in English stress assignment. It is also recognised that onset-less syllables are cross-linguistically degenerate or marked (e.g. Kager, 1999: 99; Balogné Bérces, 2006), and it seems a reasonable hypothesis that stress assignment may prefer licit over degenerate syllables. In English, independent evidence for the tendency for onset-less syllables to reject stress comes from Scottish surnames (Heinz Giegerich, personal communication): the presence or absence of an onset can account for the stress contrasts seen in *Mć.**In**tosh* versus *Mc.**Kín**ley*, and *Mć.**A**fee~Mc̀**A**fée* versus *Mc.**Háf**fie*. Although the presence or absence of onsets for word-initial syllables cannot predict secondary stress behaviour, the sensitivity of stress assignment to onsets does appear to be a significant tendency among the monomorphemic and bound-root base #LLL words.

Given the exceptions to the onset argument, it may be the case that there are other phonological factors determining the second-syllable placement of secondary stress in #LLL words. One possibility suggested here is the role of vowel height. There are languages where stress is sensitive to vowel height (Hayes, 1995: 297), and Rice (1996) proposes just such an analysis for primary stress in English. Rice's

[5] Words which are consistently #LLH have been excluded.

argument is designed to account for the apparently exceptional penultimate stress of American English *Alabáma*, *cadáver*, *erráta*, *Candelábra*, *Buchánan*, where the vowel of the penultimate syllable of these words is phonologically short /æ/. Rice argues that this low vowel, while phonologically short, is phonetically longer than other vowels and diphthongs which are phonologically long, e.g. /iː, uː, eɪ/, and that, consequently, the behaviour of /æ/ may be special with respect to stress. When we consider that lowering of vowel height corresponds to an increase in sonority (Giegerich, 1992: 133), the argument that vowel height may influence stress placement again seems plausible.[6]

The vowel height argument, like the argument for onset-less syllables, seems to capture a tendency in the data. For all of the words with consistent second-syllable stress, the vowel of the second syllable is lower than that of the initial syllable, as shown in table 5:

Word	Vowel quality	
	1st syllable	2nd syllable
Byèlorússia	/ɪ/	/ɛ/
Bucàramánga	/ʊ/	/æ/
egàlitárian	/ɪ/	/æ/
Ekàterínburg	/ɪ/	/æ/
Epàminóndas	/ɛ/ or /ɪ/	/æ/
Schehèrazáde	/ɪ/	/ɛ/

Table 5: vowel height as a predictor of secondary stress placement

In contrast, for the consistently initial-stressed #LLL words in table 2, the initial vowel is nearly always lower and never higher than the vowel in the second syllable,

[6] A case in point is that low sonority syllabic segments may be ignored by stress assignment completely in word-final position, hence the ante-penultimate primary stress in *intérminab**le*** (/l/) and *rélevanc**y*** (/y/).

e.g.: *Mèditerránean* (/ɛ/ versus /ɪ/); *Torremolinos* (/ɒ/ versus /ɪ/). And, for three out of the four words with variable secondary stress, the vowels of the first and second syllables are of the same height, making it plausible that secondary stress should fall on either syllable: *apotheosis* (/æ/ versus /ɒ/);[7] *Iphigeneia* (/ɪ/ in both syllables); *Navratilova* (/æ/ in both syllables). However, as with the argument for the effect of word-initial onset-less syllables, vowel height cannot be the whole story as far the assignment of secondary stress to #LLL words is concerned: variable secondary stress occurs in *episiotomy*, but with initial-syllable /ɛ/ and second-syllable /ɪ/, secondary stress would only be expected on the initial syllable.

In conclusion, although there do appear phonological tendencies with respect to the placement of second-syllable secondary stress in #LLL monomorphemic and bound-root base words, these tendencies suffer exceptions and cannot, therefore, predict with absolute accuracy which syllable will receive secondary stress. Given this indeterminacy, and the majority status of the initial-stressed pattern with pre-tonic #LLL, it will henceforth be proposed that anything other than consistent initial-syllable secondary stress in #LLL words has exceptional status. This is also the stance taken by Pater (1995: f.n. 11): Pater proposes that instances of second-syllable secondary stress in #LLL monomorphemic words are lexically marked to override the default *àbracadábra*. The proposal that second syllable stress in words like *Epàminóndas* is exceptional, and particularly the proposal that this stress is lexically marked, will be seen to have important implications for the theoretical modelling of second-syllable foot-head preservation in chapters 7 and 8.

Before moving on, the behaviour of a couple of exceptional #LHL monomorphemic and bound-root base words must be discussed. It was noted in §5.2.1 that variable secondary stress occurs in *Amontillado*, and consistent initial-syllable secondary stress in *apotheosis*; both defy the hypothesis for monomorphemic and bound-base words made in chapter 3, which was that these words should have consistent second-syllable secondary stress, as in #LHL *Mon**òn**gahéla*. It is possible that the variable secondary stress in *Amontillado* is not

---

[7] There is arguably a small difference in height between/æ/ and /ɒ/, but the height contrast here is much smaller than found in consistent second-syllable stress words like *Bucàramánga* (/ʊ/ versus /æ/).

an exception to phonological generalisations if the potential for sonorant codas to coalesce with preceding syllable nuclei (or alternatively reduced vowel nuclei to function as weightless) (§2.3.2) is taken into account. With coalescence – *a*[mn̩]*tillado* – the pre-tonic syllable weight sequence would be #LLL, not #LHL, so that initial-syllable secondary stress would actually be predicted by the phonology (cf. #LLL *àbracadábra*). And without coalescence – *a*[mɒn]*tillado* – second-syllable secondary stress would be predicted. In sum, variable secondary stress in *Amontillado* may be due to variable syllabification. However, an explanation is not apparent for the second exception to the #LHL generalisation – *àpotheósis*.

5.3 Support for second-syllable foot-head preservation

We can now compare the secondary stress behaviour of words which are candidates for second-syllable foot-head preservation (§5.1), and those which are not (§5.2), to establish the strength of the evidence for second-syllable foot-head preservation in light-initial words. This comparison will only consider words with an #LLL pre-tonic sequence. As noted in §5.1, very few examples of #LLH words were found which are candidates for second-syllable preservation, and so no meaningful comparison with monomorphemic and bound-root base words can be carried out. #LHL sequences will also not be examined: it was confirmed in §5.2 that phonology alone tends to predict second-syllable secondary stress with #LHL (e.g. *Monòngahéla*) the great majority of the time – little contrastive evidence for second-syllable preservation will be available. #LHH words are not considered: both monomorphemic and bound-root base words, and words which are candidates for second-syllable preservation, rarely have this pre-tonic syllable-weight sequence (§5.1 and §5.2).

To try and ensure that all the words analysed here do indeed have light first, second and third syllables, and are therefore incontrovertibly #LLL, some particularly controversial cases of syllable weight have been factored out: words with orthographic geminates or /sC/ clusters following their first, second or third syllables; words where the weight of the third syllable varies between heavy and light (e.g. *machic**o**lation*, with [əʊ~ə] in its third syllable); words containing a prefix attached to a non-cranberry root where this prefix is consonant-final; and, finally, words with a consistently reduced vowel in their first or second syllables. The controversy concerning both orthographic geminates and /sC/ clusters with respect to syllable weight was discussed in §4.4; the potentially special status of consistently reduced vowels with respect to stress assignment was also discussed there. The reason for excluding words with consonant-final prefixes attached to non-cranberry roots was discussed in §5.1.2: Giegerich (1999) proposes that such prefixes do not syllabify with the following word in the stratum one phonology, therefore potentially rendering the initial syllables of words like *inoculation* (base [*ocul*-]$_R$, also in *ocular*) or *inàmoráto* (base [*amor*-]$_R$, also in *amorous*) heavy, not light.

For Wells (2000), a total of 61 #LLL words which were candidates for second-syllable foot-head preservation were examined (appendix E). Out of these 61 words, 2 had variable secondary stress; the rest had secondary stress consistently upon their second syllables (i.e. in preserving position). A total of 16 monomorphemic and bound-root base #LLL words were examined for Wells, of which 4 had consistent second-syllable secondary stress, and the rest variable or consistent initial-syllable secondary stress. For Jones (2003), 51 #LLL words which were candidates for second-syllable preservation were examined, 1 with variable secondary stress, 2 with consistent initial-syllable secondary stress, and the rest with consistent second-syllable stress (appendix F). 15 monomorphemic and bound-root base words were examined from Jones, of which 4 had consistent second-syllable stress, and the rest variable or consistent initial-syllable secondary stress.

Chi-square analyses were performed to compare the behaviour of #LLL words which were candidates for second-syllable foot-head preservation, and monomorphemic and bound-root base #LLL words. The results of these analyses for Wells (2000) and Jones (2003) are given in (2) and (3) respectively. The results of the chi-square analyses were highly significant in both cases: $\chi^2 = 43.830$, 1 df, $p<.001$ for Wells (2000); $\chi^2 = 27.075$, 1 df, $p<.001$ for Jones (2003). However, in both cases, there was the same problem as experienced with the chi-square analyses in chapter 4: expected frequencies below 5. Happily, there were highly significant results with both Fisher's exact test and the likelihood ratio for both Wells and Jones. All of the statistical results, along with the chi-square contingency tables, are given in (2) and (3).

(2) Second-syllable foot-head preservation in #LLL words from Wells (2000)

Wells (2000)		Candidate for second-syllable preservation?	
		Complex, second-syllable preservation expected	Monomorphemic and bound-base
Secondary stress	Variable or consistent initial-syllable stress	2 (11.1 expected) 3.3%	12 (2.9 expected) 75.0%
	Consistent second-syllable stress	59 (49.9 expected) 96.7%	4 (13.1 expected) 25.0%

χ^2 = 43.830, 1 df, p<.001. One expected frequency is less than 5.

Fisher's exact test: p<.001 (2- tailed)

Likelihood ratio = 37.418, 1df, p<.001

(3) Second-syllable foot-head preservation in #LLL words from Jones (2003)

Jones (2003)		Candidate for second-syllable preservation?	
		Complex, second-syllable preservation expected	Monomorphemic and bound-base
Secondary stress	Variable or consistent initial-syllable stress	3 (10.0 expected) 5.9%	10 (3.0 expected) 66.7%
	Consistent second-syllable stress	48 (41.0 expected) 94.1%	5 (12.0 expected) 33.3%

χ^2 = 27.075, 1 df, p<.001. One expected frequency is less than 5.
Fisher's exact test: p<.001 (2- tailed)
Likelihood ratio = 23.580, 1df, p<.001

In conclusion, we have highly significant statistical support for the existence of second-syllable foot-head preservation in #LLL words. Neither the existence of monomorphemic #LLL exceptions like *Ep**à**minóndas* or *N**à**vratilóva~Navr**à**tilóva,* nor the occurrence of preservation failure in examples like *m**ì**scegenátion* (*misc**é**genate*), negates the very existence of second-syllable foot-head preservation. However, as we will see in chapter 8, both the presence of monomorphemic exceptions, and the occurrence of preservation failure, does have important implications for the precise theoretical handling of English weak stress preservation.

5.4 Initial-syllable foot-head preservation

It was noted in §3.4.1 that the evidence for initial-syllable preservation in #LLL and #LLH words is anticipated to be less overwhelming: initial-syllable preservation, e.g. ***cà**pitalístic* (***cá**pitalist*), will result in a stress pattern that is independently the default in #LLL words, e.g. monomorphemic *àbracadábra*. #LLL and #LLH words which are candidates for initial-syllable preservation are examined here. As hypothesised in §3.4.1, there are virtually no #LH- words which are candidates for initial-syllable preservation (just *melodramatic*, which is variably #LHL and #LLL).

Only one example of apparent initial-syllable preservation failure was found in all of the #LLL and #LLH words from either pronouncing dictionary: just *T**rì**politánia~Trip**ò**litánia* (*T**rí**poli*) from Jones (2003), out of a total of 56 #LLL words that do not have consistent schwa in their second syllable.[8] Initial-syllable preservation is apparently uniformly successful in the data from Wells (2000) – all words which are candidates for initial-syllable preservation have consistent initial-syllable secondary stress. The lists of all #LLL and #LLH words from both Wells and Jones are given on appendices I and J.

#LLL words which were candidates for initial-syllable preservation were compared to #LLL monomorphemic and bound-base words in order to find statistical evidence for initial-syllable preservation. Words which were sometimes or always #LLH were not examined on the grounds that the data was ambiguous: for most of the #LLH words which were candidates for initial-syllable preservation, the weight of the third syllable varies between heavy and light, for which variable lenition of the vowel of the embedded *-ize* suffix, from [aɪ] to [ɪ] (e.g. *stèri**li**zátion*), is largely responsible.

As with the statistical analysis for second-syllable preservation in §5.3, certain steps were taken to try and control syllable weight as much as possible: words with orthographic geminates or /sC/ clusters following their first, second or third syllables were removed, as were words with consistent schwa in their second

[8] Given the initial-syllable default stress pattern seen in *àbracadábra*, it is odd that there should be any initial-syllable preservation failure, as in *Trip**ò**litánia*: even in the absence of preservation, we might expect phonology to reinforce the initial-syllable secondary stress pattern. Interestingly, however, given the discussion in §5.2.2, the vowel of the second syllable of *Tripolitania* is lower than that of the initial syllable, so that we might expect the second syllable to attract stress at the expense of the first.

syllable. Prefixes do not figure in the data sets, so no steps needed to be taken with respect to prefixation.

A total of 40 #LLL words where initial-syllable preservation was expected were examined from Jones (2003), one of which one has variable secondary stress. A total of 36 words which were candidates for initial-syllable preservation were examined from Wells (2000), all of which appeared to display consistent initial-syllable preservation. A total of 16 monomorphemic and bound-root base words from Wells (2000) were examined, of which 8 had variable or consistent second-syllable secondary stress, and 8 had consistent initial-syllable secondary stress. For Jones (2003), a total of 15 monomorphemic or bound-root base words were examined, 6 of which had consistent initial-syllable stress, and 9 of which had variable or consistent second-syllable stress.

The behaviour of #LLL words which were candidates for initial-syllable preservation and #LLL monomorphemic and bound-root base words were compared in chi-square analyses. For both Wells (2000) and Jones (2003), the results were highly significant: $\chi^2 = 21.273$, 1 df, $p<.001$ for Wells; $\chi^2 = 24.246$, 1 df, $p<.001$ for Jones. Again, these results were unreliable, but both Fisher's exact test and Likelihood ratios were highly significant. All of the statistical results, along with the chi-square contingency tables, are given in (4) and (5).

(4) Initial-syllable foot-head preservation in #LLL words from Wells (2000)

Wells (2000)		Candidate for initial-syllable preservation?	
		Complex, initial-syllable preservation expected	Monomorphemic and bound-root base
Secondary stress	Variable or consistent second-syllable stress	0 (5.5) 0.0%	8 (2.5) 50.0%
	Consistent initial-syllable stress	36 (30.5) 100.0%	8 (13.5) 50.0%

χ^2 = 21.273, 1 df, p<.001. One expected frequency below 5.

Fisher's exact test: p<.001 (2-tailed)

Likelihood ratio = 22.469, 1 df, p<.001

(5) Initial-syllable foot-head preservation in #LLL words from Jones (2003)

Jones (2003)		Candidate for initial-syllable preservation?	
		Complex, initial-syllable preservation expected	Monomorphemic and bound-root base
Secondary stress	Variable or consistent second-syllable stress	1 (7.3 expected) 2.5%	9 (2.7 expected) 60.0%
	Consistent initial-syllable stress	39 (32.7 expected) 97.5%	6 (12.3 expected) 40.0%

χ^2 = 24.246, 1 df, $p<.001$. One expected frequency below 5.

Fisher's exact test: $p<.001$ (2-tailed)

Likelihood ratio = 22.612, 1 df, $p<.001$

We therefore have highly significant statistical support for the existence of initial-syllable foot-head preservation in #LLL words.

5.5 Other types of left-edge stress preservation

So far in this chapter and the previous one, canonical left-edge stress preservation has been dealt with: preservation of primary stress from a single, immediately embedded word, where the embedded word has just a single stress pattern. Other types of left-edge stress preservation are dealt with in this section, specifically:

(i) Preservation where the embedded word has a variable stress pattern (§5.5.1).

(ii) Preservation of lower levels of stress (§5.5.1 also).

(iii) Preservation from words which are more-deeply embedded in the embedding word (§5.5.2): e.g. *ícon* → *iconoclastic*, rather than *icónoclast* → *iconoclastic*.

A note on why these issues are interesting and/or important is needed.

First of all, with respect to (i), it is important to consider whether the stress of an embedded word is variable or not: if it is, then variable secondary stress in the embedding word is not an indication of preservation failure, as has been the case so far, but rather indicates successful preservation of the embedded word's stress behaviour.

With respect to (ii), while discussions of stress preservation deal with the preservation of the primary stress of the embedded word, a case for the preservation of other levels of metrical structure (stresslessness, secondary stress) was presented in §3.3.1: the argument that initial-syllable secondary stress of words like *àcademícian* is due to preservation of initial-syllable secondary stress and/or second-syllable stresslessness from embedded words like ***àca****démic*.

Finally, with respect to (iii), the possibility of stress being preserved from more-deeply embedded words is extremely important. The phonological cycle predicts that stress is inherited only from immediately embedded words: *originality* may preserve stress from *oríginal*, but may not preserve stress from *órigin* (see chapters 1, 6, 7 and 8 for relevant discussion). However, it must at least be considered that speakers associate words more generally than is assumed by the cycle: it is plausible that a speaker might associate *órigin* and *originality*, as well as *oríginal* and *originality*. This hypothesis is argued in §6.5.4.1, where psycholinguistic theory is given some consideration. Meanwhile, relevant data is presented in §5.5.2: examples of putative stress preservation failure, where the 'failing' stress pattern corresponds to the stress pattern of a more-deeply embedded word, e.g. ***tòtalitárian***~*totàlitárian* (***tótal***, *totálity*).

5.5.1 Embedded words with variable stress patterns

There are a considerable number of words from Wells (2000) and Jones (2003) where the variable stress pattern of the embedded word is preserved in the embedding word.

Words where variable primary stress in the embedded word corresponds to variable secondary stress in the embedding word are given in table 6:

	Embedding word	Embedded word	Given in Wells (2000)	Given in Jones (2003)
1	**à**pplicabílity~appl**ì**cabílity	**á**pplicable~appl**í**cable	Y	Y
2	**à**ristocrátic~ar**ì**stocrátic	**á**ristocrat~ar**í**stocrat	Y	N
3	c**àr**cinogénic~carc**ì**nogénic	c**ár**cinogen~carc**í**nogen	Y	N
4	c**òm**parabílity~comp**à**rabílity	c**óm**parable~comp**á**rable	Y	N
5	c**òm**putabílity~ comp**ù**tabílity	c**óm**putable~comp**ú**table	Y	Y
6	d**è**monstrabílity~dem**òn**strabílity	d**é**monstrable~dem**ón**strable	Y	Y
7	d**è**spicabílity~desp**ì**cabílity	d**é**spicable~desp**í**cable	Y	Y
8	**ì**dealístic~id**è**alístic	**í**dealist~id**é**alist	Y	N
9	**ìn**audibílity~in**àu**dibílity	**ín**audible~in**áu**dible	N	Y
10	l**à**ryngoscópic~lar**ỳn**goscópic	l**á**ryngoscope~lar**ýn**goscope	N	Y
11	**òx**ygenátion~ox**ỳ**genátion	**óx**ygenate~ox**ý**genate	Y	N
12	ph**à**nerogámic~phan**è**rogámic	ph**á**nerogam~phan**é**rogam	N	Y
13	ph**òs**phorylátion~phosph**ò**rylátion	ph**ós**phorylate~phosph**ó**rylate	Y	N
14	pr**è**ferabílity~pref**è**rabílity	pr**é**ferable~pref**é**rable	Y	N
15	tr**àns**ferabílity~transf**è**rabílity	tr**áns**ferable~transf**é**rable	Y	Y

Table 6: variable primary stress in embedded word corresponds to variable secondary stress in embedding word [9]

For the words in table 6, it cannot be argued that the variable secondary stress in the embedding word indicates failure of left-edge stress preservation; rather, variable secondary stress in the embedding word indicates successful preservation from the embedded word.

As with the preservation of a single stress pattern, there is evidence to suggest that the preservation of a variable stress pattern can fail (table 7):

[9] Variable ***àl****veolárity~alv****è****olárity* (***ál****veolar~alv****é****olar*) has been excluded because the variable secondary stress occurs with different segmental realisations in both the embedding and embedded words.

	Embedding word	Embedded word	Given in Wells (2000)	Given in Jones (2003)
1	accl**ì**matátion	accl**í**mate~**á**cclimate	N	Y
2	ar**ì**stocrátic	ar**í**stocrat~**á**ristocrat	N	Y
3	c**à**pitalístic	c**á**pitalist~cap**í**talist	Y	Y
4	c**àr**cinogénic	c**ár**cinogen~carc**í**nogen	N	Y
5	int**è**rcalátion	**ín**tercalate~int**é**rcalate	N	Y
6	ph**à**nerogámic	ph**á**nerogam~phan**é**rogam	Y	N

Table 7: failure to preserve variable primary stress from embedded word

Although the examples in table 7 are instances of preservation failure, if, as the proportions found in the dictionaries suggest, this failure occurs the minority of the time (cp. successful preservation in table 6), then we have further evidence for left-edge stress preservation.

As noted in chapter 3, Kager (1989) suggests that words with two embedded stress variants, where only one stress variant has primary stress on the first or second syllable, may also cause variable secondary stress placement in the embedding word. Examples from the pronouncing dictionaries are given in table 8:

	Embedding word	Embedded word	Given in Wells (2000)	Given in Jones (2003)
1	**àna**stigmátic~an**à**stigmátic	**àna**stígmat~an**á**stigmat	Y	Y
2	**àrith**metícian~ar**ìth**metician	**àrith**métic$_{A,}$ ar**íth**metic$_{N}$	Y	Y
3	**èqui**librátion~eq**ùi**librátion	**èqui**líbrate~eq**úi**librate	Y	N
4	**ìnter**calátion~int**èr**calátion	**ìnter**caláte~int**ér**calate	Y	N
5	**ìnter**pellátion~int**èr**pellátion	**ìnter**péllate~int**ér**pellate	Y	Y

Table 8: possible variable preservation of secondary or absent stress

For the initial-stressed variants of embedding words like *ànastigmátic*, one may propose that there is either preservation of initial-syllable secondary stress (***à****nastígmat*) and/or preservation of second-syllable stresslessness (*àn****a****stígmat*) from the embedded word. No attempt will be made to satisfy this debate here. On the one hand, it seems plausible that there might be wholesale, rather than selective, metrical identity between words.[10] However, on the other hand, one could make the argument for it being only the salient characteristics of words – i.e. stress, rather than stresslessness – which are likely to be preserved.

Whichever elements of lower level metrical structure we argue are preserved, the initial-stressed variant ***àrith****metícian* provides reasonably good evidence for the preservation of metrical structure other than primary stress. *Àrithmetícian* has the pre-tonic syllable weight sequence #LHL, and so would be most likely to have second-syllable secondary stress in the absence of preservation, cf. monomorphemic #LHL *Mon****òn****gahéla* (§5.2). The overriding of purely phonological predictions to give ***àrith****metícian* can be accounted for if the secondary stress and/or locations of unstressed syllables in ***àrith****métic* are preserved.

With respect to the preservation of lower level metrical structure, just a single instance of failure has been found in either Wells (2000) or Jones (2003): invariant

[10] Fitting in with this hypothesis, Burzio (1994) proposes a principle of 'Metrical Consistency', which simply requires words to be as metrically consistent as possible.

***sì**nfoniétta*, from ***sì**nfonía~sinf**ó**nia*. Again, therefore, the general picture from the data is supportive of preservation.

5.5.2 Preservation from more-deeply embedded words

As will be discussed at length in chapters 7 and 8, the phonological cycle crucially predicts that phonological characteristics should only be inherited from immediately embedded words: a stress preservation relationship can only exist where the embedding and embedded words are separated by no more than one suffix. However, as is argued in the next chapter, it seems psycholinguistically plausible that speakers may associate words in a less restricted fashion, and therefore that stress could also be preserved from more-deeply embedded words: e.g. *tótal* in *totalitarian*, as well as *totálity*. This possibility has important theoretical implications: as we will see in chapters 7 and 8, other theoretical means of enforcing phonological identity between words besides the cycle – notably Output-Output Correspondence and fake cyclicity – permit identity between embedding words and more-deeply embedded words. The possibility of preservation from more-deeply embedded words is therefore considered here.

Possible instances of preservation from more-deeply embedded words are given in table 9. This list has been compiled by examining all words from both pronouncing dictionaries which were candidates for preservation and which exhibited variable secondary stress, and then sorting out those where the stress pattern of a more-deeply embedded word corresponded to the 'failing' stress variant of the embedding word.[11]

[11] The embedding words in table 9 are not included on the accompanying appendices.

	Embedding word	Embedded word	Given in Wells (2000)	Given in Jones (2003)
1	**àm**bassadórial~amb**à**ssadórial	**ám**bassy$_N$ (OED), amb**á**ssador$_N$	Y	Y
2	**àn**tipathétic~ant**ì**pathétic	**àn**tipáthic$_A$ (OED), ant**í**pathy$_N$	Y	Y
3	c**èr**tificátion~cert**ì**ficátion	c**ér**tify$_V$, cert**í**ficate$_N$	Y	Y
4	c**ò**incidéntal~co**ìn**cidéntal	c**ò**incíde$_V$, co**ín**cidence$_N$	Y	Y
5	d**ì**rectionálity~dir**èc**tionálity	d**ì**rectée$_N$ (OED), dir**éc**tional$_A$	Y	N
6	h**ù**manitárian~hum**à**nitárian	h**ú**man$_N$, hum**á**nity$_N$	Y	Y
7	**ì**conoclástic~ic**ò**noclástic	**í**con$_N$, ic**ó**noclast$_N$	Y	Y
8	**ì**sochronícity~is**ò**chronícity	**í**sochron$_N$ (OED), is**ó**chronous$_A$	Y	N
9	M**èphi**stophélean~ Meph**ì**stophélean	M**èphi**stópheles$_N$, Meph**í**sto$_N$	Y	Y
10	p**òn**tificátion~pont**ì**ficátion	p**ón**tiff$_N$, pont**í**ficate$_V$	Y	N
11	Sh**àke**speriána~Shakesp**è**riána	Sh**áke**speare$_N$, Shakesp**é**rian$_A$	Y	Y[12]
12	t**ò**talitárian~tot**à**litárian	t**ó**tal$_N$, tot**á**lity$_N$	Y	Y
13	tr**ì**angularity~tri**àn**gulárity	tr**í**angle$_N$, tri**án**gular$_A$	N	Y
14	tr**ì**angulation~tri**àn**gulátion	tr**í**angle$_N$, tri**án**gular$_A$	N	Y
15	**ù**tilitárian~ut**ì**litárian	**ú**tilise$_V$, ut**í**lity$_N$	N	Y

Table 9: preservation from non-immediate sub-constituents

[12] Strangely, this variable secondary stress is only shown in Jones (2003) with the <ia> spelling of -*iana*. The spelling with <ea> is included on Appendix B.

The variable secondary stress in all of the embedding words in table 9 cannot be taken to indicate a total failure of stress preservation failure: while preservation from the immediately embedded word does occur, this may, in fact, be the result of preservation from a more-deeply embedded word. In §6.5.4.1, frequency information is given which supports this hypothesis.

5.6 Conclusion

In this chapter, two important empirical findings have been made with respect to left-edge stress preservation:

(i) Evidence has been presented for foot-head preservation in words with light initial syllables.

(ii) Evidence has been presented which indicates that foot-head preservation is a variable, rather than consistently successful, phenomenon.

Strong statistical evidence has been presented both for initial- and second-syllable foot-head preservation. As noted in the conclusion of chapter 4, the finding of evidence for left-edge stress preservation is very important indeed: it means that phonological theory must be able to handle English weak stress preservation. Equally important is this chapter's finding that both initial- and second-syllable foot-head preservation are not consistently successful, as it was noted in chapter 4 that relative prominence preservation similarly appeared to be a variably successful phenomenon. The possibility that stress preservation can fail will be shown to have extremely important implications for the formal handling of English stress preservation in chapter 8. Meanwhile, in the next chapter, one possible explanation for the variable rather than consistent success of left-edge stress preservation is explored: word frequency effects, and their implications for morphological decomposability.

There have been some additional interesting findings in this chapter. First of all, we saw in §5.1 that, in Jones (2003), relative prominence preservation is significantly more associated with failure than second-syllable foot-head preservation. Possible explanations for this are considered in §6.5.1 and §9.3.2; a means of formalising this difference in behaviour is also proposed in chapter 9. The data presented for non-canonical left-edge preservation in §5.5 also offered some food for thought. In §5.5.1, evidence was presented which indicated that the variable stress patterns of embedded words may be preserved in their embedding words, and that levels of metrical structure below that of main stress – secondary stress and stresslessness – may be preserved in embedding words. These examples provided further evidence for weak stress preservation. An observation which will acquire some importance in the following chapters of the thesis is that presented in §5.5.2: the possibility that more-deeply embedded words influence the stress of embedding words. In chapter 6, it is shown that this hypothesis is reasonable in light of hypotheses about the nature of lexical access and observations about word frequency. Subsequently, in §8.3.2, vowel shortening data is presented which strongly supports the hypothesis that phonological characteristics may be inherited from more-deeply embedded words. These combined arguments for preservation from more-deeply embedded words will be seen to have serious implications for the phonological cycle.

Chapter six: word frequency and left-edge stress preservation

6.0 Introduction

In this chapter, it is shown that word frequency is one factor which can account for the variable success of left-edge stress preservation that was recorded in chapters 4 and 5. Psycholinguistic work on word frequency shows that embedded words can be perceived as being present inside their embedding words to a greater or lesser degree. From this, it is hypothesised here that it may not be enough for a speaker to know a word from which stress may be preserved; rather, this embedded word additionally has to be frequent enough to be reliably perceived, and so for stress to be preserved from it. This hypothesis is supported by the findings presented in this chapter.

Three disclaimers are necessary with respect to the data analysis presented in this chapter. First, with the exception of the analysis in §6.5.1, the quantitative analyses presented in this chapter are restricted to words which are candidates for relative prominence preservation (e.g. *ămb**à**ssadórial* ← *àmb**á**ssador* – see chapters 3 and 4). This is due to the insufficient sizes of some of the categories for preservation in light-initial words (chapter 5) that are candidates for foot-head preservation alone. There were instances of what appeared to be failure of foot-head preservation alone in light-initial words, e.g. ***mì**scegenátion~misc**è**genátion* (*misc**é**genate*) in chapter 5, but, with so few examples of preservation failure where the pre-tonic syllable weight sequence is incontrovertibly #LLL, there is little from which a statistical analysis could determine a cause of preservation failure.[1] The statistical analysis in §6.5.1 does, however, indicate that the same types of frequency effects hold over second-syllable foot-head and relative prominence preservation, and this is reflected in the theoretical analyses proposed in subsequent chapters.

[1] It is vital that the pre-tonic sequence is #LLL when examining foot-head preservation: #H- sequences are candidates for relative prominence preservation; #LH- sequences may have second-syllable secondary stress independent of preservation (cf. *Mon**ò**ngahéla*). In chapter 5, there were just three #LLL words with evidence of failure of second-syllable foot-head preservation once controversial cases of syllable weight (those involving /sC/ clusters, orthographic geminates and non-cranberry prefixations) were removed: *Boliviano*, *miscegenation* and *vaticination*. There was just one example of initial-syllable foot-head preservation failure: *Trìpolotánia~Tripòlitánia* (*Trípoli*), regardless of any controls for syllable weight.

Second, a three-way distinction between consistently successful relative prominence preservation, variably successful relative prominence preservation, and consistently unsuccessful relative prominence preservation is not utilised in the analyses in this chapter. Instead, consistently successful relative prominence preservation is classed as ‘preserving stress,’ and both variably successful and consistently unsuccessful relative prominence preservation are classed under a single category of ‘non-preserving stress’. This collapsing of categories is, again, due to limitations in the amount of data available for analysis: there are just six words with consistent initial-syllable secondary stress, but it was not desirable to exclude these six words given the sample sizes required for some of the statistical analyses employed in this chapter. The collapsing of variably and consistently failing relative prominence preservation into a single category seems reasonable: both variable and consistent initial-syllable secondary stress are hypothesised to be genuine sub-types of relative prominence preservation failure.

Third, only data for relative prominence preservation from Jones (2003) is examined (appendix B). This is because the number of cases in the non-preserving category for relative prominence preservation in Jones (2003) is much larger than that in Wells (2000), meaning Jones gives much more to go on when trying to establish a cause for relative prominence preservation failure. A possible explanation for the disparity in proportions between Wells and Jones is explored in §6.5.2.

The chapter is structured as follows. In §6.1, a frequency analysis of relative prominence preservation is carried out which is comparable in its method to Hammond’s (2003a) analysis of pre-tonic preservation (§2.5.1). This analysis finds a different relationship between word frequency and preservation to that argued for by Hammond, and a possible explanation for the conflicting results is presented in §6.1.1. In §6.2, it is shown that the relationship found to exist between word frequency and relative prominence preservation is psycholinguistically plausible; Hay’s (2003) hypothesis concerning relative frequency is introduced. In §6.3 and §6.4, data analysis based on Hay’s relative frequency hypothesis is carried out. The empirical plausibility of the results presented in this chapter is defended in §6.5, before the chapter is concluded in §6.6.

6.1 Comparison with Hammond's analysis of pre-tonic preservation

In chapter 2, we saw that the quality of pre-tonic vowels in words like *cònd*[ɛ]*nsátion* may be preserved from their embedded words (*cond*[ɛ́]*nse*), but that this type of preservation can also fail, as in *ìnf*[ə]*rmátion* (*inf*[ɔ́ː]*rm*). In §2.5.1, we saw that Hammond (2003a) reports an effect of cumulative frequency upon pre-tonic vowel quality preservation which is argued to account for its variable success: preservation is more likely to fail the more frequent the embedded word, as well as the more frequent the embedding word. Here, an investigation largely comparable in method to Hammond's is carried out in order to see whether a similar effect of cumulative frequency holds for relative prominence preservation; it will be seen that it does not.

As in Hammond's (2003a) analysis, only words with the *-ation* (compound) suffix are examined. In total, there were 98 *-ation* words from Jones (2003) where relative prominence preservation was expected: 68 words displayed consistent relative prominence preservation ('preserving stress'), and 30 words displayed relative prominence preservation variably or never ('non-preserving stress'). Hammond's (2003a) frequency counts are the total number of occurrences of a word in the Brown Corpus, and are approximately per million.[2] Here, frequency counts from the British National Corpus (Leech, Rayson & Wilson, 2001) ('BNC') have been used: the frequency counts used are also per million, but, unlike the Brown Corpus, the BNC frequencies have the advantage of distinguishing between nouns and verbs which have the same orthographic form, e.g. *coordinate*$_{\text{N}}$ versus *coordinate*$_{\text{V}}$.[3] The frequencies are lemma frequencies: they include frequencies of inflected forms. Where a word is not listed in the BNC, it has been recorded as having a frequency of zero. All 98 embedding *-ation* words, together with their BNC frequencies and the frequencies of their embedded words, are given in appendix K.

The scatterplots in figures (1a) and (1b) show raw embedding BNC frequency plotted against raw embedded BNC frequency for words from Jones (2003); as noted, these are words where relative prominence preservation is expected. Each

---

[2] The Brown Corpus contains approximately 1,026,604 words (Hammond, 2003a: f.n. 12).

[3] Hammond (2003a: 8) has to throw out words like *pigmentation* because, in his frequency information, he cannot separate counts for *pigmént*$_{\text{V}}$ and *pígment*$_{\text{N}}$ (*-ation* nouns are deverbal, not denominal). Having to throw out data is problematic unless one has an excess of it, and there is no excess of data here.

marker stands for a single case – an embedding -*ation* word where preservation is expected. The shape of the marker indicates the embedding word's preservation behaviour. For each case, the frequency of the embedding word is shown on the horizontal axis, and the frequency of its embedded word (from which preservation is expected) is shown on the vertical axis. Figure (1a) includes only 96/98 of the -*ation* words examined: *consideration* and *continuation* are excluded to aid readability of the graph, as the frequencies of their embedded words are extremely high (289/million for *consider*, and 283/million for *continue*). From (1a), it is clearly obvious that the great majority of -*ation* embedding words and their embedded words have frequencies below 20/million (90/98 of -*ation* words in all): therefore, just words with frequencies below 20/million are shown in figure (1b) so that their distribution can be seen in more detail.[4] In figure (1b), it is clear that cases in the non-preserving category cluster along the bottom of the scatterplot, while preserving cases tend to occur higher up on the graph; the significance of this observation will become apparent later in the chapter.

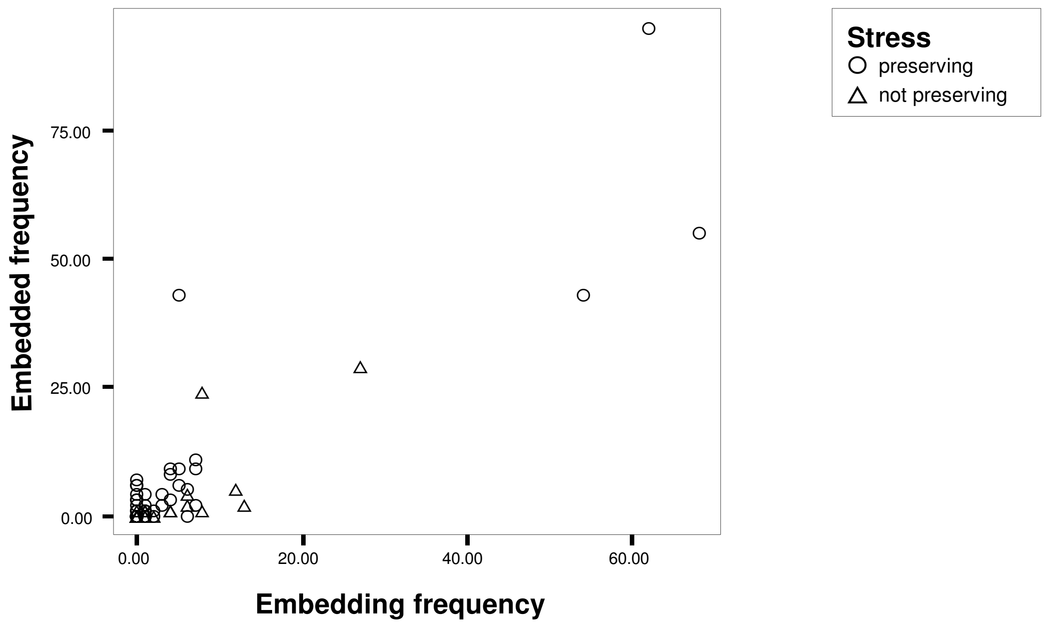


Figure (1a): Effect of embedded and embedding frequency upon relative prominence preservation in -*ation* words (BNC frequencies <100/million)

[4] Neither axis in figure (1b) goes up to 20/million because no word has this exact frequency for its embedding or embedded word – frequencies fall above or below it.

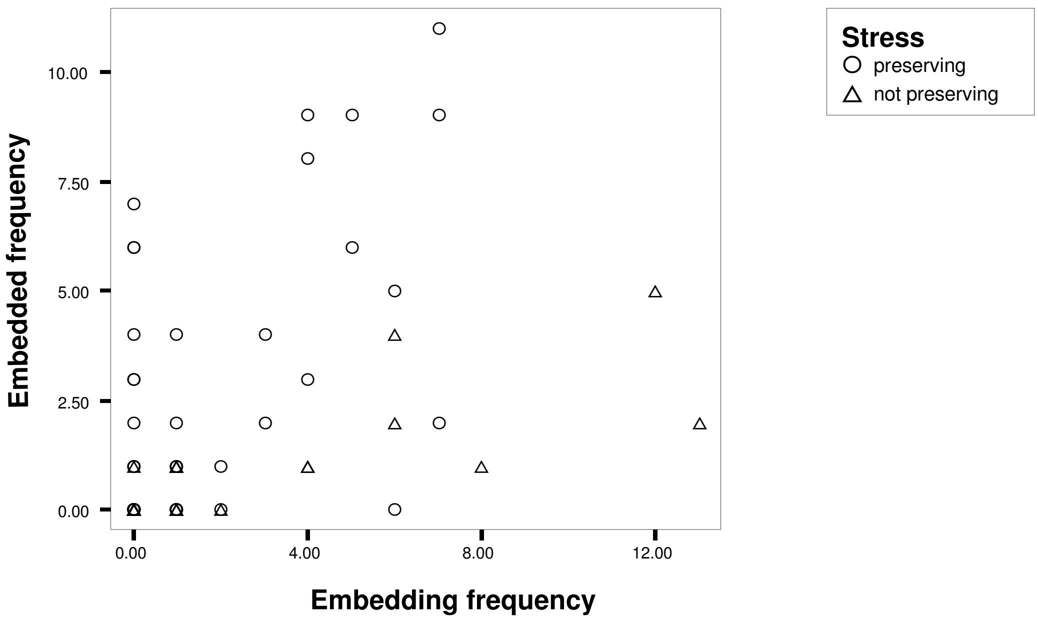


Figure (1b): Effect of embedded and embedding frequency upon relative prominence preservation in *-ation* words (BNC frequencies <20/million)

Following Hammond (2003a), regression analysis was performed on the raw BNC frequencies to determine the nature of the relationship between frequency and preservation. However, rather than using the linear regression method employed by Hammond, binary logistic regression (forced entry) was used. This is because the outcome variable – whether there is consistent relative prominence preservation or not – is categorical. The category 'preserving stress' was coded as a 0, and that of 'non-preserving stress' as a 1. A total of 8 examples where embedding or embedded frequency were higher than 20/million were excluded – logistic regression can be sensitive to outliers.

The results of this first regression analysis are given in table 1:

Model
χ^2 = 14.813, 2df, p=.001 -2LL = 96.784 R^2 = .152 (Cox & Snell); R^2= .214 (Nagelkerke) Tolerance = .744, VIF = 1.345
Embedding BNC frequency
Wald statistic significance = p<.01, 1df Exp *b* = 1.474, 95% C.I. 1.105 (Lower) to 1.965 (Upper) B = 0.388 (S.E. 0.147)
Embedded BNC frequency
Wald statistic significance = p<.025, 1df Exp *b* = 0.550, 95% C.I. 0.338 (Lower) to 0.894 (Upper) B = -0.599 (S.E. 0.248)

Table 1: logistic regression for -*ation* words with BNC raw frequencies <20/million

Overall, there is a highly significant effect of word frequency upon relative prominence preservation behaviour (model χ^2 = 14.813, 2df, p=.001). Both embedding and embedded frequency have significant effects on stress preservation behaviour (Wald statistic significance p<.01 and p<.025 for embedding and embedded frequency respectively). There is a positive relationship between embedding frequency and preservation failure: the greater the embedding frequency, the more likely relative prominence preservation is to fail (Exp $b > 1$; 95% C.I. does not cross 1). This matches Hammond's (2003a) observation for pre-tonic preservation: it too is more likely to fail with increased embedding frequency. Here, there is a negative relationship between embedded frequency and preservation failure: the greater the embedded frequency, the less likely relative prominence preservation is to fail (Exp $b < 1$; 95% C.I. does not cross 1). The effect of embedded frequency is therefore opposite of what is found by Hammond (2003a) for pre-tonic vowel-quality preservation: Hammond reports a cumulative effect, so that increased frequency of the embedded word makes preservation more likely to fail, not less as is found here.

Hammond's (2003a) analysis of pre-tonic vowel reduction only deals with raw frequency. However, to include the 8 words with one or both frequencies above 20/million, regression analysis was also performed using log-transformed frequencies (log-transformation reduces the impact of outliers). The use of log-transformed frequencies is also motivated for another reason: logarithmic frequencies better resemble how "humans process frequency information" (Hay & Baayen, 2002: 208). Hay & Baayen argue that, in human perception of frequency information, "the difference between 10 and 20 is more important/salient than the difference between 1010 and 1020" (Hay & Baayen, 2002: 208). Log-transformation alters the raw data in such a way that it better resembles the proportional differences perceived by humans.[5] So that the log of zero frequencies could be taken, 1 was added to each raw frequency.

[5] Following Hay & Baayen, all logarithmic transformations in this chapter are to the base of the natural logarithm.

Model
χ^2 = 8.761, 2df, p<.025 -2LL = 111.967 R^2 = .086 (Cox & Snell); R^2 = .121 (Nagelkerke) Tolerance = .390, VIF = 2.566
Log embedding BNC frequency
Wald statistic significance = p<.025, 1 df Exp *b* = 2.535, 95% C.I. 1.153 (Lower) to 5.573 (Higher) B = 0.930 (S.E. 0.402)
Log embedded BNC frequency
Wald statistic significance = p<.025, 1 df Exp *b* = 0.363, 95% C.I. 0.167 (Lower) to 0.788 (Higher) B = -1.013 (S.E. 0.395)

Table 2: logistic regression for *-ation* words: log-transformed BNC frequencies

There is an overall significant effect of log frequency upon relative prominence preservation behaviour (χ^2 = 8.761, 2df, p<.025). Both log embedding and log embedded frequency are individually significant predictors of stress preservation behaviour (Wald statistic significance is p<.025 for both embedding and embedded frequency). As log embedding frequency increases, non-preserving stress becomes more likely (Exp *b* > 1; 95% C.I. does not cross 1). As before, the opposite effect was detected for log embedded frequency: the more frequent the embedded word, the less likely non-preserving stress is (Exp *b* < 1; 95% C.I. does not cross 1).

In conclusion, with respect to the frequency of the embedded word, a different frequency effect has been found to that reported for pre-tonic vowel quality preservation by Hammond (2003a). Hammond argues for an effect of cumulative frequency: preservation is more likely to fail as the frequencies of both embedding and embedded words increase. Here, preservation has been shown to be more likely to fail with increased frequency of the embedding word, but decreased, not increased, frequency of the embedded word. A possible explanation for this conflict of outcomes is now given in §6.1.1.

6.1.1 Reassessment of Hammond's results

Hammond's (2003a) argument for an effect of cumulative frequency upon pre-tonic preservation appears to be psycholinguistically plausible: Hammond points out that psycholinguistic research like Taft (1979) has reported cumulative frequency effects. However, it is odd that two such similar phenomena – preservation of vowel quality versus stress, both in stratum one vocabulary – are associated with opposite effects with respect to the frequency of the embedded word.

As it turns out, Hammond's (2003a) argument for an effect of cumulative frequency is not as convincing as one would hope. This is because Hammond (2003a) misuses the R^2 statistic with respect to determining the direction of correlation:

> There was a significant correlation with the frequency of the derived word: R^2 = .3441, p = .0002. There was also a significant correlation with frequency of the base form: R^2 = .1363, p = .0267 […] The correlation with the frequency of the derived form means that the more frequent the form is, the more likely it is to undergo reduction. The correlation with the frequency of the base form means that the more frequent the base form, the more likely the derived form is to undergo reduction (Hammond, 2003a: 9-10).

The problem with Hammond's statement here is that the R^2 results, which he cites as indicative of positive or negative correlation, can only tell us **how much** of the outcome that the predictors account for – they tell us nothing about the **direction** of correlation (this would be the job of the correlation coefficient, *r*, in the linear regression analysis Hammond uses).[6] Of course, Hammond may have based his observations about the direction of correlation on other statistics that are not included in Hammond (2003a), but this is left to the reader's good faith.

In light of the ambiguity in Hammond's reporting, a statistical analysis of Hammond's data was carried out to obtain information about the directions of correlation for both embedding and embedded frequency. Hammond (2003a)

[6] There is this same ambiguity in reporting in Hammond (2003b) and in its corresponding paper, Hammond (2004: 359). Data from Hammond (2003a) is used in the following analysis because it is the largest data set provided by Hammond.

commendably makes his work replicable by including a list of all 36 words he analyses, along with their frequencies and those of their embedded words. Hammond also reports the preservation behaviour of these words (based on vowel qualities shown in the MRC Psycholinguistic Database). As the outcome is categorical – a full vowel or a reduced vowel – Hammond's data was analysed using logistic regression.[7] The majority of embedding words have a full pre-tonic vowel, and so this behaviour was coded with a 0; reduced vowels were coded as a 1. Logistic regression analyses were performed on both raw and log-transformed frequencies, the results of which are given in tables (3a) and (3b):

Model
χ^2 = 20.259, 2df, p<.001 -2LL = 24.057 R^2 = .430 (Cox & Snell); R^2 = .608 (Nagelkerke) Tolerance = .771, VIF = 1.297
Log embedding frequency
Wald statistic significance = p<.05, 1 df Exp *b* =1.949, 95% C.I. 1.152 (Lower) to 3.297 (Higher) B = .667 (S.E. 0.268)
Log embedded frequency
Wald statistic significance = p<1, 1 df Exp *b* = 1.013, 95% C.I. 0.909 (Lower) to 1.129 (Higher) B = .013 (S.E. 0.055)

Table (3a): reanalysis of Hammond's data (raw frequencies)

[7] The results in tables (3a) and (3b) differ from those reported in Hammond (2003a) because, whereas logistic regression is used here, Michael Hammond (personal communication) uses regular linear regression. It is potentially problematic to use linear regression when there is a categorical outcome: linear regression absolutely assumes the relationship between variables is linear, but, when the outcome is categorical, the relationship isn't generally linear (Field, 2005: 220).

Model
χ^2 = 20.950, 2df, p<.001 -2LL = 23.366 R^2 = .441 (Cox & Snell); R^2 = .623 (Nagelkerke) Tolerance = .710, VIF = 1.409
Log embedding frequency
Wald statistic significance = p<.01, 1 df Exp *b* = 9.322, 95% C.I. 1.865 (Lower) to 46.596 (Higher) B = 2.232 (S.E. 0.821)
Log embedded frequency
Wald statistic significance = p<1, 1 df Exp *b* = 1.264, 95% C.I. 0.501 (Lower) to 3.185 (Higher) B = .234 (S.E. 0.472)

Table (3b): reanalysis of Hammond's data (log-transformed frequencies)

The regression analyses in (3) do show that frequency can significantly predict pre-tonic preservation behaviour (p<.001 for the model χ^2 in both cases). The logistic regression models also show that there is a positive relationship between (log) embedding frequency and preservation: the more frequent the embedding word, the more likely pre-tonic preservation is to fail. However, in both models, there is a big problem when we come to assess the individual effect of embedded frequency: although Exp *b* is above 1 in both (3a) and (3b), suggesting that there is a positive relationship between embedded frequency and vowel reduction, as argued by Hammond, this result is not reliable – the 95% Confidence Interval for Exp *b* crosses 1 in both (3a) and (3b). Although the regression model indicates that preservation is more likely to fail when the more frequent the embedded word – the cumulative frequency effect reported by Hammond – this result is by no means reliable.

In sum, sound evidence is still needed in order to support the argument that there is an effect of cumulative frequency upon pre-tonic preservation.[8] If such

[8] A reliable result may be obtained with a larger data set. One possible source of additional words is those which Hammond had to exclude because the frequency of the verbal base could not be isolated

evidence is found, then it will need to be considered how two apparently antithetical frequency effects can exist alongside one another. A speculative solution is given in the conclusion of the thesis.

6.2 Frequency, lexical access and preservation

6.2.1 The dual-route model and preservation

In §6.1, it was shown that word frequency significantly affects the success of relative prominence preservation. This relationship is anticipated in light of the effects of word frequency upon lexical access in the dual-route model.

There are models of lexical access where either the route of full morphological decomposition (e.g. Taft, 1985 [Hay, 2003: 7]) or of direct whole-word retrieval (e.g. Butterworth, 1983 [Hay, 2003: 7]) is employed for accessing affixed words. Under the route of full morphological decomposition, a morphologically complex word is processed by decomposing it into its constituent morphemes, and then looking up the lexical entries associated with these individual morphemes. Under whole-word retrieval, a lexical entry corresponding to the whole word is looked up. As its name implies, the dual-route model of lexical access proposes that both morphological decomposition and whole-word access may play a role in accessing morphologically complex words.

In the dual-route model, the decomposed route may be used to access a complex word by its morphological constituents. However, a characteristic of the dual-route model that will become crucial in accounting for stress preservation failure is that, even for transparently complex words, the whole-word route of lexical access is always available: complex words have their own lexical representations in the speaker's memory, allowing a complex word to be looked up directly without reference to its morphological constituents (Hay, 2003: 7).

Whole-word storage of complex words as proposed by the dual-route model does not economise on memory, but does reduce any computational effort involved in accessing complex words; indeed, humans do not seem to treat memory as a "scarce resource" (Pinker, 1999: 153). And, because the decomposed route is nevertheless available for regular complex words in the dual-route model, a speaker

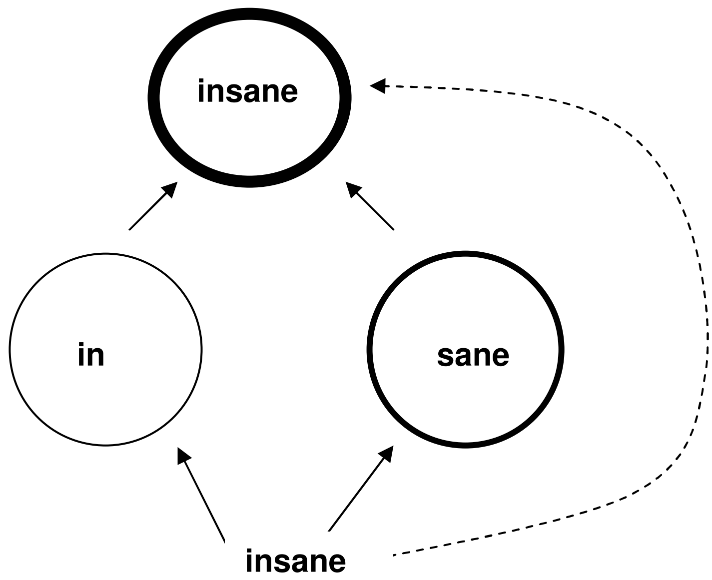
---

(see f.n. 3). This problem could be solved by using frequency information that distinguishes nouns from verbs, e.g. the BNC, used here in §6.1.

can still be sure to access a word, even if his or her memory of the whole word is weak, or if he or she has never even encountered the word before. As pointed out by McQueen & Cutler (1998: 423), the dual-route model appears to be the most successful: neither the whole-word route nor the decomposition route can, by itself, account for empirical observations. McQueen & Cutler give evidence which shows that morphological decomposition does occur in some instances, but that it is an unsuitable proposal in other situations. For example, in the case of pseudo-prefixation (e.g. pseudo-prefixed ***mis****ery* versus prefixed ***mis****place*), it would be highly uneconomical if decomposition was mandatory – there would be a processing cost from mis-parsing pseudo-prefixed forms (McQueen & Cutler, 1998: 416-7). However, decomposition of prefixed forms does seem to occur sometimes: for example, in Taft & Forster's (1975) lexical decision task, non-words took longer to be rejected if they were from a bound root (e.g. *vive* in *revive*) than if they were from a pseudo-root (e.g. *lish* in *relish*) (McQueen & Cutler, 1998: 414-5), indicating an awareness of the morphological constituents of prefixed forms.

A dual-route model of lexical access is adopted here: this is the model of lexical access in which Hay (2001, 2003) couches her hypotheses about relative frequency, and the role of relative frequency effects in relative prominence preservation is explored at length as the chapter progresses. However, Hay (2003: 7) notes that the relative frequency hypothesis is not crucially reliant upon specifically the dual-route model of lexical access: any model of lexical access will do, as long as it allows speakers to perceive morphological complexity as a gradient characteristic. It will also be assumed here that the two routes are in direct competition, as in the 'race' models of the dual-route literature (e.g. Baayen, 1992; Fraunfelder & Schreuder, 1992). Under the race model, in any given attempt to access a word, either the decomposed route will access a word's meaning first, or the whole-word route will – the two routes compete against one another, in parallel. Not all dual-route models assume a race (e.g. Caramazza et al., 1988), but the race model has been shown to be very successful at modelling human performance (Anshen and Aronoff, 1988: 647; Schreuder & Baayen, 1995 [Pinker, 1999: 155], among others). Finally, I assume that, in any single instance of access of a complex word, either the whole-word or decomposed route wins outright. It is likely that this is overly simplistic (cf. Baayen & Schreuder, 2000), but it allows for simpler argumentation in the coming chapters.

A schematised illustration of the dual-route model is given in figure 2:



The solid line indicates the decomposed route. The dashed line indicates the direct route. The width of a circle's outline indicates the resting activation of the lexical entry or 'node'. Direct access is more likely here, as *insane* has a higher resting-activation level than *sane* or *in-*.

Figure 2: Schematised dual route model from Hay (2001: 1045; 2003: 11)

As shown in figure 2, for a morphologically complex word where both the direct and decomposed routes of access are available, which of the routes wins depends upon the strengths of the memories of the lexical entries involved – specifically, their 'resting-activation' levels. There is a direct relationship between frequency and the resting-activation level of a lexical entry. Frequency information is automatically and unconsciously recorded in the memory, so that every exposure to a form increases the strength of its representation in the memory. The stronger the memory of a form, the closer its resting level is to the threshold at which the lexical entry is recognised in lexical access – akin to leaving a television set on standby, as opposed to turning it off at the wall. Consequently, lexical entries with higher resting activation levels are recognised more quickly. In figure 2, the resting-activation level of *insane* is higher than that of *sane*, meaning that access via the whole-word route will be quicker than access via the decomposed route. (As we will see later in the chapter, the resting-activation level of the affix also plays a rôle in determining the route of lexical access.)

The method of lexical access – the decomposed route or the whole-word-route – affects the mental representations of morphologically complex words. Vitally, decomposed access increases the strength of the relationships between the embedding word's lexical entry and those of its embedded morphological constituents (Hay, 2003: 8): decomposition reinforces the morphological complexity of the embedding word. Conversely, the whole-word route does not reinforce the links between the embedding word's lexical entry and those of its morphological constituents. The consequences of repeated whole-word access are that:

> Words which are more prone to whole-word access appear less affixed, undergo semantic drift, proliferate in meaning, and are implemented differently in the phonetics. They are effectively free to become phonologically and semantically liberated from their bases and acquire idiosyncrasies of their own (Hay, 2003: 16).

Token word frequency indicates the resting-activation levels of lexical entries: the more times a speaker processes a word, the higher the resting-activation level of its lexical entry will be. If an embedding word is of a high frequency, then it will have a high resting-activation, and will be more likely to be accessed directly than via the decomposed route. For this reason, high absolute frequency is characteristic of embedding words which have strayed from their bases in their semantics and phonology (Bybee, 1985: 118; Hay, 2001: 1042 and references cited therein).

Given that high frequency of the embedding word weakens the relationships between embedding words and their embedded morphological constituents, it makes sense that preserving stress was shown to be less likely with increased frequency of the embedding word in §6.1. Preserved stress indicates a relationship between an embedding word and its embedded word through phonological similarity (Cutler, 1980, 1981; Kenstowicz, 1996: 370): preserving ***ămbà****ssadórial* has the same stress contour over its first two syllables as ***àmbá**ssador*; the same cannot be said for non-preserving ***àmba**ssadórial*. Because preservation is an indicator of a morphological relationship, if this morphological relationship weakens, then so should the likelihood of preservation. And one factor expected to weaken the relationship between the embedding and embedded word is increased frequency of the embedding word.

We have, therefore, an explanation for why increased frequency of the embedding word should increase the likelihood of non-preserving stress. Increased frequency of the embedding word makes direct, whole-word access of the embedding word more likely; in turn, repeated whole-word access of the embedding word weakens the relationship between the embedding word and the embedded word – the embedded word is not being referred to in lexical access of the embedding word. As the embedding word drifts from its embedded word, its morphological structure becomes opaque, resulting in the embedding word's phonology conforming to that of a monomorphemic word – stress preservation fails. Without frequent reference to the embedded word, and thus the embedded word's phonology, in lexical access, the embedding word 'forgets' the stress pattern of the embedded word.

It was also noted in §6.1 that there is an effect of the frequency of the embedded word: the less frequent the embedded word, the more likely relative prominence preservation is to fail. A possible explanation for this is now considered in §6.2.2, in the form of Hay's (2001, 2003) argument for relative frequency effects.

6.2.2 Hay (2001, 2003) and relative frequency

Hay (2001, 2003) argues that the effects of frequency are not absolute: it is not automatically the case that embedding words with a high absolute frequency will be accessed directly, and low frequency embedding words accessed via the decomposed route. Hay proposes that which route of lexical access is chosen depends upon the ratio of the frequency of the embedding word to the frequency of the embedded word: an embedding word must be frequent enough **relative** to the frequency of its embedded word to be accessed via the direct route. The implication of Hay's argument for relative frequency is that, even if an embedding word is of a low frequency, it may still not be accessed by morphological decomposition if it is more frequent than its even lower frequency embedded word. This contrasts with Bybee's argument, which proposes that only high frequency embedding forms should be accessed via the direct route.

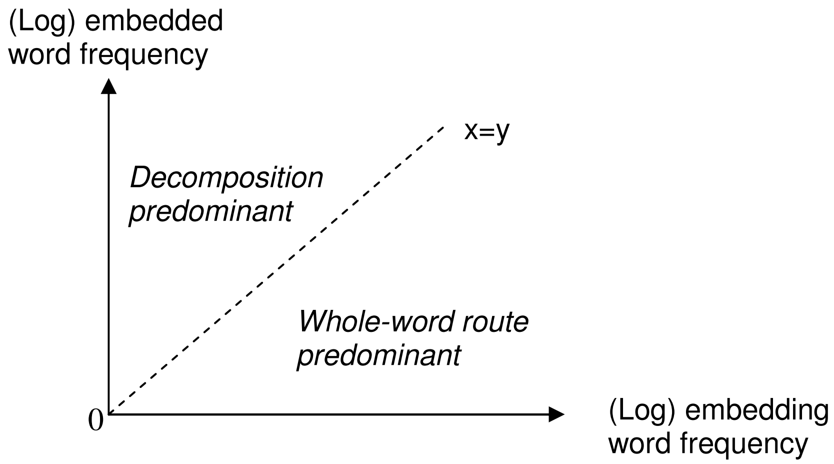
One set of experiments which support Hay's relative frequency hypothesis is her study of semantic drift. Hay shows that high frequency embedding words are no more likely to drift from their embedded words' semantics than low frequency embedding words; however, embedding words which are more frequent than their embedded words are more likely to under go semantic drift than embedding words which are less frequent than their embedded words (Hay, 2001, 2003: 110-114). Relative frequency also sheds much light upon the topic of morphological productivity (Hay & Baayen, 2002, 2003).

The frequency results from §6.1 are not observations about word-specific relative frequency effects, but rather observations about absolute frequency. Nevertheless, these results do indicate that the data may be amenable to a relative frequency analysis. We observed in §6.1 that preservation failure becomes more likely with increased frequency of the embedding word. Although Hay argues against a simple effect of absolute frequency of the embedding word, she does show that absolute frequency effects may be symptomatic of relative frequency effects: high frequency embedding words are more likely to be more frequent than their bases than low frequency embedding words (Hay, 2003: 101). The effect of embedding frequency reported in §6.1 may therefore be indicative of a relative frequency effect. It is also pertinent that a negative relationship between the frequency of the embedded word and preservation failure was noted in §6.1: the statistical analyses indicated that preservation becomes less likely the less frequent the embedded word is, regardless of the frequency of the embedding word. In sum, the relationship between preservation failure and frequency should be explored with respect to word-specific relative frequency effects.

Further particulars of the relative frequency analysis are presented in §6.2.2.1 and §6.2.2.2. The first relative frequency analysis of relative prominence preservation is then presented in §6.3.

6.2.2.1 The parsing line

Under Hay's relative frequency hypothesis, the point at which an embedding word is frequent enough, relative to its embedded word, to be accessed via the whole-word route is the 'parsing line'. In Hay (2001, 2003), the parsing line is defined as the point where the frequency of the embedded word equals the frequency of the embedding word. This is shown in figure 3:



The line x=y is the parsing line. Words falling above the line are more prone to access via the decomposed route: the frequency of the embedded word is higher than the frequency of the embedding word. Words falling below the line are more prone to direct access: the embedding word is more frequent than the embedded word.

Figure 3: the x=y parsing line (Hay, 2001, 2003)

The x=y parsing line proposed by Hay (2001, 2003), although very intuitive, is without independent empirical justification. The exact ratio of relative frequency at which decomposed and whole-word access are equally likely to occur has since been refined by Hay & Baayen (2002). (Hay & Baayen (2002) follows Hay (2003) in real time.) Hay & Baayen (2002) use a psycholinguistic model for morphological parsing, Matcheck (Baayen, Schreuder & Sproat, 2000; Baayen & Schreuder, 2000), to empirically establish the location of the parsing line. Matcheck can be used to model parallel-race lexical access (Hay & Baayen, 2002: 212): it gives individual times at which the whole-word and decomposed forms are available. Whichever route completes in the shortest time is likely to be the access route used by a speaker. (Validation for Matcheck as a predictor of speakers' lexical-access routes is given in Hay & Baayen (2002).) Hay & Baayen's investigation shows a bias in favour of whole-word access: when embedding and embedded words are of equal frequency, the whole-word route is preferred; indeed, the whole-word route may be preferred even when the embedding word is less frequent than the embedded word.[9] Hay & Baayen argue that this result is empirically plausible:

> [T]he x=y division in effect weighs up the effort involved in retrieving the base against the effort in retrieving the derived form. It does not take into consideration the added task of retrieving the affix and any subsequent calculations which may be associated with parsing, both of which add to the effort involved in successfully decomposing a word into its parts (Hay & Baayen, 2002: 217).

Hay & Baayen (2002) give an exact location for the revised parsing line. As noted in §6.3 below, the location of this is such that it does not immediately appear to be appropriate for the data investigated in this chapter. In §6.3, I argue that this may not, in fact, be the case, for reasons that are explored further in §9.3.[10] For now, I will work with the idea that there will be a higher ratio of embedding word frequency to embedded word frequency in words accessed by the whole-word route than in words accessed by decomposition, and so a higher ratio of embedding to embedded word frequency in words which are classed as having non-preserving stress. This is the approach taken in §6.3.

It is important to note that the parsing line is not assumed to be an absolute: complex words are not assumed to be always and fully decomposed because they fall above the parsing line, and, equally, are not assumed to be always and fully accessed holistically if they fall below the parsing line. Hay (2003) and Hay & Baayen (2002) argue that morphological decomposition is a continuum, with both decomposed and

---

[9] Hay & Baayen's revised parsing line was established by running English bimorphemic words suffixed with *-ness* through Matcheck. Hay & Baayen ran the same test for several other affixes, and showed that although the line's position varied depending upon the length of the suffix, the same bias towards whole-word parsing was apparent in each case.

[10] It appears to be the case that the words Hay & Baayen use are generally of a much higher frequency than the words investigated here. When the CELEX data for relative prominence preservation was plotted on a scatterplot and Hay & Baayen's revised parsing line applied to it, virtually all of the plots fell below it – the parsing line appeared to be too high (§6.3)

whole-word access playing a role in the access of most words; relative frequency effects determine how much each method of access contributes for any given complex word. The parsing line is the dividing point at which the likelihood of whole-word access and decomposition are equal; the further a complex word is from this line, the more one method of access dominates for this particular form.

6.2.2.2 Phonotactics and decomposition

Word frequency is not the only factor which affects morphological decomposition; Hay (2003) argues that morpheme-juncture phonotactics also play a role.

In line with work on probabilistic phonotactics and speech perception (e.g. McQueen, 1998; van der Lugt, 1999), Hay proposes that morpheme-juncture phonotactics help a speaker break a word down into its constituent morphemes on a probabilistic basis: if the phonotactics at a morpheme boundary are unlikely to occur morpheme-internally, the unlikely phonological transition is a good cue to a speaker to recognise the morpheme boundary. For example, a speaker is likely to recognise that *inhumane* is prefixed, as /nh/ is very unlikely to occur morpheme-internally in English; in contrast, the /ns/ transition in *insincere* is seen morpheme-internally in *fancy* and *tinsel*, making *insincere* less prone to morphological decomposition than *inhumane* (Hay, 2003: 15-16). Hay (2003) focuses upon the probabilities of different affixes, and shows that phonotactics is used to segment both nonce and real words.

For virtually all of the embedding words examined in this chapter, the phonotactics between the embedded word and the suffix are a high probability CV transition, rather than a lower probability CC transition. This follows from the fact that the suffixes involved are the vowel-initial suffixes -*ation*, -*ion* and -*ity*, not consonant-initial suffixes: when suffixes are vowel-initial, no consonant clusters will be created by adding a vowel-initial suffix to a consonant-final base; rather, any base-final consonant will form an onset in the embedding form, e.g. *insénsitive* → *in.sèn.si.tí.vi.ty*.[11] (In support of this argument, Hay & Baayen (2003: 121) show that vowel-initial -*ation* and -*ity* tend to create phonotactically legal junctures that do not

[11] CV transitions are argued to be characteristic of stratum one forms in Lexical Phonology (Raffelsiefen, 1999; Hay, 2003: 159-60). Both Hay and Raffelsiefen argue that the phonological differences between stratum one and stratum two forms negate any need the strata themselves. However, in chapter 7, it is shown that strata are required in order to model phonological opacity effects.

facilitate morphological segmentation.)[12] Although there is the possibility that not all CV transitions are of absolutely equal probability morpheme-internally (e.g. /dɪ/ may have a different probability to /tuː/), the overall high probability of CV transitions indicates that phonotactics is unlikely to be a fruitful area to research in the preservation data or to interfere with the frequency results much.[13] Phonotactics are therefore not considered any further in the analyses of relative prominence preservation.

6.3 Relative frequency and *-ation*

In this section, it is shown that the *-ation* data from §6.1 behaves in a way compatible with Hay's (2003) and Hay & Baayen's (2002) observations about relative frequency.

To get the most out of the data set, and also make the data comparable to Hay's, frequencies were collected from the CELEX lexical database (Appendix L). CELEX (Baayen et al., 1995) contains frequencies based on the number of times a word occurs in the 17.9 million-token Birmingham COBUILD corpus (Renouf, 1987), making the frequencies considerably more sensitive than the per-million BNC frequency counts used in §6.1. As with the BNC frequency counts, the CELEX frequencies used are lemma frequencies.

Following Hay & Baayen, the CELEX frequencies were log-transformed as $\log_e x$. The CELEX frequencies for the *-ation* words are plotted on the scatterplot in figure 4 below; the plot markers indicate whether the embedding word displays preserving or non-preserving stress (as in §6.1, the latter category is a grouping of

[12] Very interestingly, Hay & Baayen show that the affixes they examine with high probability junctural phonotactics (i.e. affixes that are not likely to be parsed out) also tend to create "high frequency forms, which have low frequency base forms" (2003: 118). It is therefore very plausible that stress preservation failure should occur in level one words that are characterised by high probability junctural phonotactics.

[13] A very small number of words have a VV transition at the juncture with *-ation*, and it is possible that this would have a different general probability than CV. These examples have not been removed in the analyses presented in this chapter, as I felt the benefit of this was outweighed by the benefit of using the largest data set possible.

both variable and consistent initial-syllable secondary stress). Linear regression lines are shown for both preserving and non-preserving stress.[14]

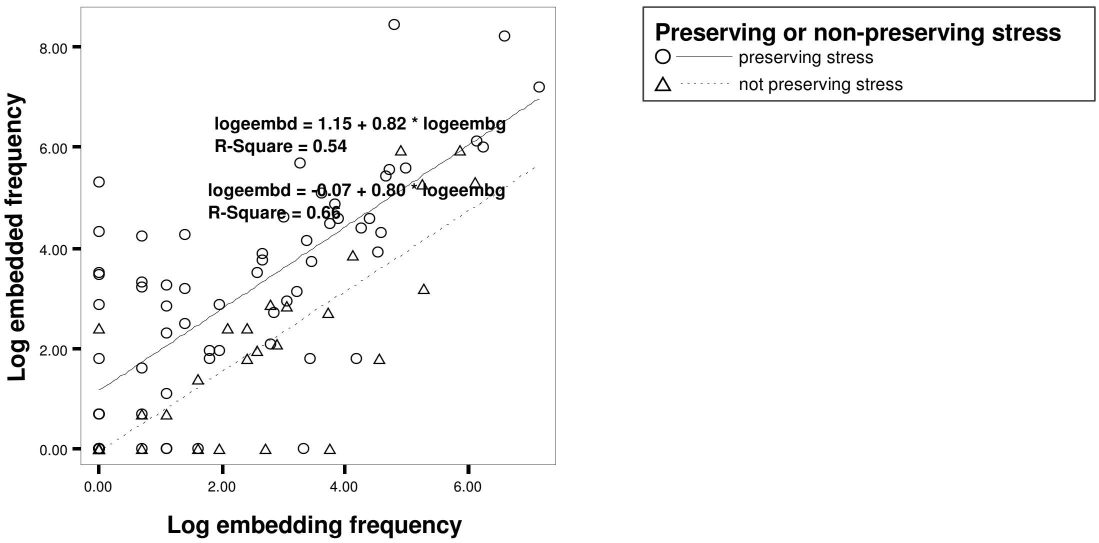


Figure 4: Frequency and stress-preservation behaviour for *-ation*

In figure 4, the linear regression lines have very similar positive slopes (.80 versus .82), but the regression line for non-preserving stress is lower than the regression line for preserving stress (the y-intercept for non-preserving stress is -0.07, compared to 1.15 for consistently preserving stress). The height of the regression line is important in determining the pattern of relative frequency in a data set, as explained by Hay & Baayen:

> [A] high intercept reflects an overall pattern in which base frequencies tend to be high relative to derived frequencies. That is, it reflects a distribution in which many words are prone to parsing, and very few are prone to whole word access. A low intercept, on the other hand, would reflect a distribution in which a larger proportion of forms fall below the parsing line. Such a distribution has a larger proportion of forms which are prone to whole word access (Hay & Baayen, 2002: 222).

Therefore, the higher position of the regression line for preserving than non-preserving stress in figure 4 indicates that words in the non-preserving category are

[14] The regression lines were calculated by the graph function in SPSS. There is an overall significant relationship between log embedding and log embedded frequency: see the discussion of collinearity in footnote 16, below.

more likely to have been accessed via the whole-word route than words in the preserving category. This is exactly in keeping with our hypothesis: it was argued in §6.2 that access via the whole-word route will increase the probability of relative prominence preservation failure, because whole-word access weakens the relationship between the embedding words and their embedded constituents.

Hay & Baayen (2002: 210) point out that the slope of the regression line will also reflect the overall pattern of relative frequency in a group of data: the steeper the regression line is for a particular category, the fewer the number of points in that category which are likely to fall below the parsing line (a steeper regression line is more likely to cross the parsing line). As the slopes for both preserving and non-preserving stress for -*ation* are about the same in figure 4 (.82 versus .80), the main difference between preserving and non-preserving stress here is the position of the y-intercept.

As discussed in §6.2.2.1, Hay & Baayen (2002) revise Hay's (2003) proposal for an x=y parsing line. This revised parsing line has an intercept of 3.76 and a positive slope of 0.76, and so would fall above the regression line for **preserving** stress in figure 4 – not what we would expect. However, it is not necessarily the case that Hay & Baayen's revised parsing line is incompatible with the data analysed here. Under the fake cyclic analysis of weak stress preservation proposed in chapter 8, whole-word access will not automatically result in preservation failure in the word in question. Fake cyclicity proposes that an embedding word is stored along with its stress pattern, and that this stress pattern will still be stored even under access via the whole-word route. Access via the whole-word route will make the stress pattern of the embedding word vulnerable to regularisation as a result of lexically gradual change; however, this change will not occur immediately, if ever. It is therefore possible that, as in figure 4, the regression line for words with preserving stress will fall below Hay and Baayen's revised parsing line, indicating that they are accessed via the whole-word route – words may be accessed via the whole-word route yet still have a preserving stress pattern. In §9.3, I show, in detail, how stress preservation failure is not simply attributable to frequency effects, but that some types of fake cyclic preservation will be more prone to failure than others, independent of frequency.

Taking, for now, Hay's earlier, purely intuitive approximation of x=y, we get a highly significant result between relative frequency and relative prominence preservation behaviour for *-ation* (based on CELEX frequency counts) in a chi-square analysis: $\chi^2 = 13.499$, 1df, p<.001. The chi-square contingency table is given in (1).

(1) Relative frequency and relative prominence preservation in *-ation* words[15]

		Relative frequency	
		Embedding > embedded frequency	Embedded > embedding frequency
Stress	Non-preserving stress	16 (8.9 expected) 51.6%	6 (13.1 expected) 13.0%
	Preserving stress	15 (22.1 expected) 48.4%	40 (32.9 expected) 87.0%

$\chi^2 = 13.499$, 1df, p<.001

In conclusion, the initial analysis presented in this section indicates that relative frequency is very likely to play a role in whether or not *-ation* words consistently preserve the relative prominence contour of their base.

6.3.1 Logistic regression analysis with CELEX frequencies for *-ation*

For the sake of completeness, the logistic regression analysis carried out in §6.1 with BNC frequencies was repeated with the log CELEX frequencies for *-ation* words. The results are given in table 4.

[15] 21 words were excluded where the frequency of the embedding and embedded words were equal.

Model
χ^2 = 17. 676, 2df, p<.001 -2LL = 103.053 R^2 = .165 (Cox & Snell); R^2= .233 (Nagelkerke) Tolerance = .484; VIF = 2.067[16]
Log embedding frequency
Wald statistic significance = p<.005, 1 df Exp *b* = 2.060, 95% C.I. = 1.295 (Lower) to 3.277 (Higher) B = 0.723 (S.E. 0.237)
Log embedded frequency
Wald statistic significance = p= .001, 1 df Exp *b* = .463, 95% C.I. = 0.299 (Lower) to 0.717 (Higher). B = -0.770 (S.E. 0.223)

Table 4: logistic regression for *-ation* words (log-transformed CELEX frequencies)

Overall, the model based on CELEX word frequency is a very significant predictor of relative prominence preservation behaviour (χ^2 = 17. 676, 2df, p<.001). Both log embedding and log embedded CELEX frequency were individually significant predictors (Wald statistic significance p<.005 for log embedding frequency and p= .001 for log embedded frequency). As log embedding frequency increased, so did the likelihood of non-preserving stress (Exp *b*>1, 95% C.I. does not cross 1). As log embedded frequency increased, non-preserving stress became less likely (Exp *b*<1, 95% C.I. does not cross 1). In sum, the same effects of embedding and embedded

[16] There is likely to be some degree of collinearity between log embedding frequency and log embedded frequency. Hay (2003) and Hay & Baayen (2002) show, that for many suffixes, there is a significant relationship between embedding ('derived') and embedded ('base') frequency. There is such a correlation between the CELEX log embedding and log embedded frequencies for *-ation* with both parametric and non-parametric correlation tests: Pearson = .718, p<.01; Spearman's rho = .668, p<.01 (in both 1- and 2-tailed tests).

Collinearity makes it harder to assess the individual contributions of independent variables to the outcome of the model. However, as both Tolerance and VIF (measures of collinearity) for the regression models given in this chapter fall below critical levels, the regression results may be argued to still merit consideration.

frequency have been found for -*ation* over two sets of word frequency information: that from the BNC, and now that from CELEX.

6.4 Individual suffixes

A factor not explored yet in this chapter is the role individual suffixes play upon morphological decomposition. Hay & Baayen (2002) show that affixes can vary in their decomposability as a consequence of word-specific relative frequency effects. This variation occurs among affixes because different affixes vary in how many of the words containing the affix fall above or below the parsing line. If most of the words containing a particular affix are prone to decomposed access as a result of word-specific relative frequency effects, then the affix is likely to have an independent lexical representation with a reasonably high resting-activation level (e.g. -*ness*), and any form containing this affix is more likely to be decomposed. Conversely, if most of the words containing an affix have relative frequencies such that they are prone to whole-word access, the affix itself will have a low resting-activation level (Hay, 2003: 16), and any forms containing this affix will be less prone to decomposed access.

The resting-activation levels of individual affixes are in addition to any relative frequency effects. If we take two words which are in approximately the same relative frequency ratio with their base, but whose affixes have different resting-activation levels, then these words are not equally likely to be decomposed. Hay (2003) gives the example of -*ish* versus -*ic*:

> *Grayish* and *scenic*, for example, have roughly similar frequency profiles (per 17.4 million: *gray*: 32, *grayish*: 1542, *scene*: 30, *scenic*: 1995). However, because the affix in the first word has a higher resting activation level than the affix in the second word, we expect *grayish* to be more decomposable than *scenic* (Hay, 2003: 157).

In light of the possible variation in the decomposability of different suffixes, the data analyses presented so far in this chapter must be further refined. So far, -*ation* has been analysed. Words ending in -*ation* are the result of two separate processes: addition of -*ion* to embedded words ending in -*ate* (e.g. *anticipate*, *anticipat**ion***), as well as affixation with -*ation* in one go (e.g. *deforest*, *deforest**ation***) (cf. Marvin, 2002: 65). For safety, this variable should be factored out of the data set.

Words created by addition of -*ion* to -*ate* words constitute the vast majority of -*ation* words examined so far; statistical tests have therefore been repeated with just these words, the results of which are given in §6.4.1. Another suffix, -*ity*, is also examined, in order to see whether the effects of frequency upon relative prominence preservation are evident with more than one suffix.

6.4.1 Analysis of -*ion*

For -*ion*, effects of both absolute and relative frequency were examined for.

A total of 12 embedding words were removed which consisted of -*ation* attached to the embedded word in a single step, e.g. *deforest*, *deforest**ation***. (These removed words are indicated by shading in appendix L.) A total of 86 words consisting of -*ion* added to an -*ate* base, e.g. *anticipate*, *anticipat**ion***, remained.

In order to test for an effect of absolute frequency, logistic regression was performed (re: §6.1, above). Binary logistic regression (forced entry) was performed upon the log-transformed CELEX frequencies, the results of which are given in table 5:

Model
χ^2 = 8.851, 2df, p<.025 -2LL = 94.826 R^2=.098 (Cox & Snell); R^2= .140 (Nagelkerke) Tolerance = .430; VIF = 2.324
Log embedding frequency
Wald statistic significance p<.025, 1df Exp *b* = 1.789, 95% C.I. 1.089 (Lower) to 2.939 (Higher) B = 0.582 (S.E. 0.253)
Log embedded frequency
Wald statistic significance = p<.01, 1df Exp *b* = .537, 95% C.I. 0.341 (Lower) to 0.848 (Higher) B = -0.621 (S.E. 0.233)

Table 5: Logistic regression for -*ion* words (log-transformed CELEX frequencies)

Overall, the model constructed upon log CELEX frequency was a significant predictor of relative prominence behaviour in *-ion* words ($\chi^2 = 8.851$, 2df, $p<.025$). Log embedding and log embedded frequency were individually significant predictors (Wald statistic significance $p<.025$ and $p<.01$ for log embedding and log embedded frequency respectively). Non-preserving stress became more likely as the log frequency of the embedding word increased (Exp $b>1$, 95% C.I. does not cross 1). Non-preserving stress became less likely as the log frequency of the embedded word increased (Exp $b<1$, 95% C.I. does not cross 1).

Word-specific relative frequency effects also hold in the analysis of *-ion*. The scatterplot in figure 5 shows log embedding and log embedded frequency plotted against one another; the relative prominence preservation behaviour of the embedding word is indicated by the marker.

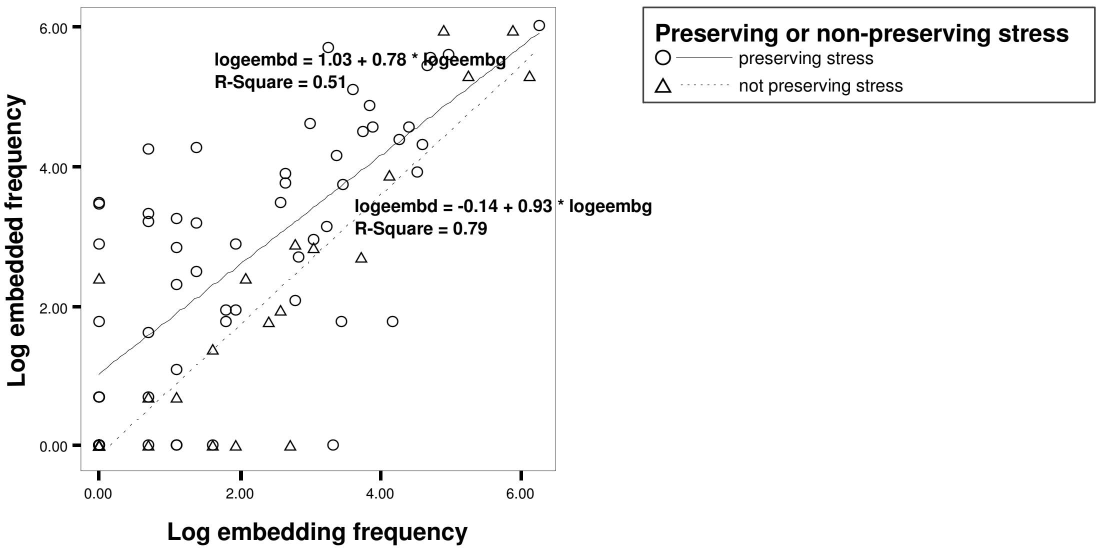


Figure 5: frequency and preservation behaviour for *-ion*

The linear regression line for non-preserving stress has a lower y-intercept than that for words with preserving stress. This is as we would expect under the relative frequency hypothesis: a lower regression line indicates a higher ratio of embedding to embedded frequency, and therefore a greater likelihood of being accessed via the whole-word route that erodes preservation. The steeper slope of the line for non-preserving stress than for preserving stress (.93 versus .78) somewhat reduces the contrast between the two groups: a steeper regression slope means more of the points in the non-preserving category are likely to fall above the parsing line, and hence be prone to decomposed access (see §6.3, above). But, overall, the lines are in the positions we would hope, the regression line for the category 'non-preserving stress' being consistently lower on the graph than that for the category 'preserving stress'.

There is further support for the existence of a relative frequency effect in -*ion* words. When x=y is taken as an approximate parsing line, there is a significant relationship in a chi-square analysis between relative frequency and whether or not an embedding -*ion* word displayed consistently preserving stress ($\chi^2 = 9.045$, 1df, p=.005). The chi-square contingency table is given in (2):

(2) Relative frequency and relative prominence preservation in -*ion* words[17]

		Relative frequency	
		Embedding > embedded frequency	Embedded > embedding frequency
Stress	Non-preserving stress	12 (7.0) 46.2%	6 (11.0) 14.6%
	Preserving stress	14 (19.0) 53.8%	35 (30.0) 85.4%

$\chi^2 = 9.045$, 1df, p=.005

In the analysis of a single suffix, -*ion*, there is statistical support for effects of both absolute embedding and embedded frequency upon relative prominence preservation, and also a good indication that word-specific relative frequency effects occur.

[17] 19 words were excluded where embedding and embedded frequencies were equal.

6.4.2 Other suffixes

So far in this chapter, words with the compound suffix *-ation* have been analysed for frequency effects, as have a subset of these words – *-ion* suffixations. The obvious question now is whether the frequency effects recorded so far hold with any other suffix.

A significant problem with analysing another suffix is getting a suffix for which there are enough tokens for statistical analysis, particularly where non-preserving stress is concerned. After *-ion*, the suffix for which there are the most tokens among the non-preserving set is *-ity* (66 *-ity* words in total, given in appendix M: 53 words with consistently-preserving stress, 13 without).[18] The same analyses were repeated using embedding words containing *-ity* as their outermost suffix, and CELEX frequencies. No further suffixes were analysed – the quantities of tokens with non-preserving stress with any single suffix are very small indeed.

To test for an effect of absolute frequency, a logistic regression analysis was performed on the log embedding and embedded frequencies of the *-ity* data. The result was not significant (model $\chi^2 = 0.591$, 2df, $p<1$). There is insufficient data for a chi-square analysis of relative frequency for *-ity* words.[19] However, impressionistically, there is no obvious effect of relative frequency which could account for variation in the success of preservation in *-ity* words, at least with the line x=y: all of the *-ity* words have an embedded frequency higher than their embedding frequency, including those in the non-preserving category.

We need to account for why frequency can account for some of the stress preservation behaviour of *-ion* words, but for none of the stress preservation behaviour of *-ity* words.[20] To this end, the relative frequency distributions of *-ion* and *-ity* words were compared – figure 6:

[18] 66 words is too small a sample size to acceptably test for the effects of individual predictors (i.e. embedding frequency and embedded frequency) in a regression analysis (Field, 2005: 173). The regression analysis was simply performed to see if the overall model was a significant fit of the data, for which there are just enough tokens.

[19] Only 9 non-preserving *-ity* words have embedding and embedded frequencies which are not the same.

[20] Although we cannot rule out the possibility that we may find a significant frequency effect with a larger sample size of *-ity* words.

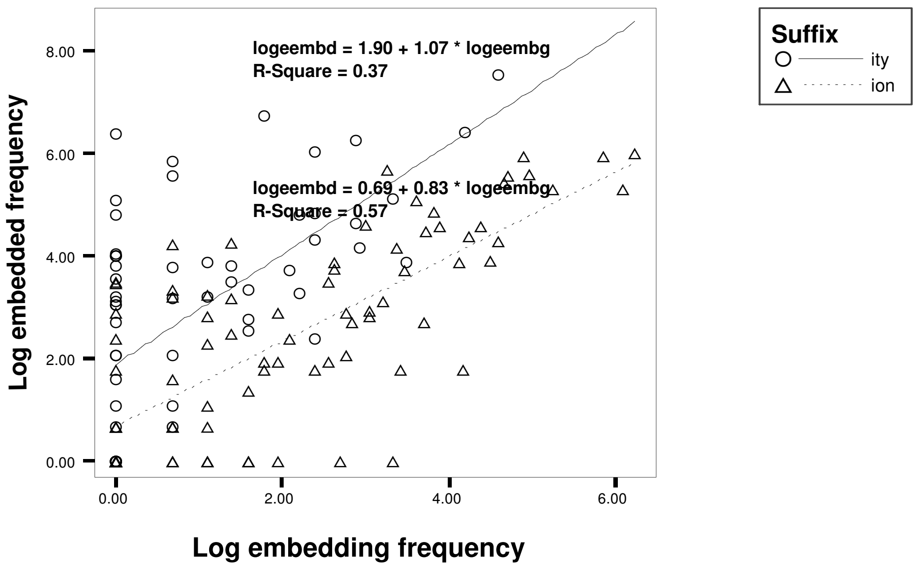


Figure 6: relative frequency distributions of -*ion* and -*ity*

The linear regression lines in figure 6 show that, with respect to frequency, -*ity* words are more likely to be decomposed – exhibit preserving stress – than -*ion* words: the linear regression line for -*ity* is higher and has a steeper slope than that for -*ion*. From this observation, it is plausible that frequency effects account for the stress preservation failure in the -*ion* group, but not the -*ity* group: in terms of frequency, -*ion* words are better disposed to preservation failure. However, this explanation is only plausible if there is something else aside from frequency causing the stress preservation failure in the -*ity* group. Happily, it is indeed possible that this is the case: the regression analyses given so far in this chapter for -*ation* and -*ion* show that, at best, frequency effects account for less than 25% of relative prominence preservation behaviour (this is indicated by the R^2 values); there must therefore be other variables determining the success of preservation. In §6.5.4, some of the other factors which may affect the success of relative prominence preservation are considered.

6.5 Evaluation of the frequency analyses

It has been shown that there are significant relationships between word frequency and relative prominence preservation behaviour. However, significant statistical results are meaningless unless they are also empirically plausible. The following questions must therefore be answered:

(i) Why is relative prominence preservation more prone to failure than foot-head preservation (re: §5.1.2)?

(ii) Why is there a higher proportion of relative prominence preservation failure in Jones (2003) than in Wells (2000) (re: §4.2.1)?

(iii) Why do most words in the non-preserving category have variable secondary stress, rather than consistent initial-syllable secondary stress?

(iv) Aside from frequency, what other factors contribute to stress preservation failure?

Solutions (some concrete, some speculative) to these questions are given in the following four sub-sections. All arguments indicate that the foregoing statistical analyses are empirically plausible.

6.5.1 Relative prominence versus foot-head preservation

In the chi-square analysis given in §5.1.2, we saw that relative prominence preservation appears to be significantly more prone to failure than second-syllable foot-head preservation. It will now be shown that relative prominence preservation appears to be intrinsically fallible: when frequency is held constant, words which are candidates for relative prominence preservation are less likely to have secondary stress in preserving position than words which are candidates for just second-syllable foot-head preservation. It is suggested that relative prominence preservation may be more fallible than foot-head preservation because a greater perceptual cost is associated with the failure of foot-head preservation than of just relative prominence preservation.

A logistic regression analysis was performed to see whether the type of second-syllable preservation a word is a candidate for – relative prominence or just second-syllable foot-head – has an effect upon how likely it is to exhibit successful preservation (i.e. secondary stress on its second syllable). Not all of the words included in the chi-square analysis in §5.1.2 have been examined here: this original data set contains a range of suffixes, and, in §6.4, we saw that different suffixes are decomposable to different degrees. Instead, just -*ion* suffixed words have been examined. There are a total of 86 -*ion* words which are candidates for relative prominence preservation (#HHL and #HLL words), and 69 -*ion* words which are candidates just for second-syllable foot-head preservation (#LLL words).[21] Log-transformed embedding and embedded frequencies were included as predictor variables. Which type of second-syllable preservation a word is a candidate for was also included as a categorical variable; words which are candidates for relative prominence preservation were coded with a 0, and those which are candidates for just second-syllable foot-head preservation were coded with a 1. As before, preserving stress was coded as 0, non-preserving stress as 1. A forced entry logistic regression analysis was then performed.

The results of the regression analysis are given in table 6:

[21] Raw frequencies for #LLL -*ion* words where second-syllable preservation is expected are given on appendix N.

Model
χ^2 = 28.430, 3df, p<.001 -2LL = 118.005 R^2 = .168 (Cox & Snell); R^2 = .274 (Nagelkerke)
Type of preservation
Wald statistic significance p<.001, 1df Exp *b* = 0.082, 95% C.I. 0.022 (Lower) to 0.311 (Higher) B = -2.497 (S.E. 0.678) Tolerance = .997 VIF = 1.003
Log embedding frequency
Wald statistic significance p<.05, 1df Exp *b* = 1.568, 95% C.I. 1.039 (Lower) to 2.366 (Higher) B = .450 (S.E. 0.210) Tolerance = .504 VIF = .1.985
Log embedded frequency
Wald statistic significance p<.005, 1df Exp *b* = 0.557, 95% C.I. 0.380 (Lower) to 0.817 (Higher) B = -.585 (S.E. 0.195) Tolerance = .504 VIF = 1.983

Table 6: relationship between stress preservation type and success

The logistic regression analysis shows that the type of preservation for which a word is a candidate, as well as embedding and embedded frequency, determines whether second-syllable stress preservation is successful. If a word is a candidate for relative prominence preservation by virtue of having a heavy initial syllable, it is significantly more likely to exhibit preservation failure than if it is a candidate for just foot-head preservation (i.e. it has a light initial syllable) (Wald statistic significance p<.001, 1df; Exp *b*<1, 95% C.I. does not cross 1). There were the same effects of frequency as before, indicating that the frequency effect holds over simple foot-head as well as relative prominence preservation. As embedding frequency increases, so does the likelihood of relative prominence preservation failure (Wald statistic significance p<.05, 1df; Exp *b*>1, 95% C.I. does not cross 1). As embedded frequency increases, relative prominence preservation is less likely to fail (Wald statistic significance p<.005, 1df; Exp *b*<1, 95% C.I. does not cross 1).

At first, it seems odd that relative prominence preservation would be more prone to failure than foot-head preservation: it suggests that maintaining identity of foot-head placement is more important than maintaining the identity of the relative prominences of these foot heads. In fact, Ricardo Bermúdez-Otero (personal communication) has suggested that such a situation may be perceptually well-motivated. If foot-head preservation fails, then it is not only the stress contour of the embedded word which may be lost: the segmental content of the embedded word is at risk too, as any unstressed vowel is at risk of vowel reduction: *ir*[ǽ]*scible* → **ìr*[ə]*scibílity*.[22] The cost associated with the failure of relative prominence preservation alone, with foot-head preservation being successful, is altogether less severe: the stress contour of the embedded word will be obscured by being reversed – ***àntí****cipate* → ***àntĭ****cipátion* – but, as long as foot-head preservation is successful, at least the segmental content of the embedded word is guaranteed to be visible in the embedding word. When frequency effects are set aside, perceptual differences may create a situation whereby foot-head preservation must absolutely be respected, but failure of relative prominence preservation is more tolerable. In chapter 9, it will be shown how these perceptual differences can be formalised in the stratum one constraint ranking.

6.5.2 Proportions of preservation failure in Jones (2003) vs. Wells (2000)

Before considering why there may be divergence between Jones (2003) and Wells (2000) in terms of the proportion of success of relative prominence preservation, it is worth emphasising the fact that both dictionaries exhibit **some** evidence of relative prominence preservation failure. There is, therefore, a consensus between the dictionaries that relative prominence preservation is a variable phenomenon.

A reason why there is a greater proportion of relative prominence preservation failure in Jones (2003) than Wells (2000) may be the different stress-

[22] In accordance with the argument in chapter 2, it is not assumed that all unstressed vowels will automatically reduce; however, the absence of stress obviously makes reduction a possibility.

assignment methods employed by the dictionary editors. For Jones (2003), it is likely that secondary stress patterns have been assigned on a predominantly intuitive, non-systematic basis. Jones still largely retains the stress assignments from its 1917 compilation by Daniel Jones, and Jones treated the stress of polysyllabic words as "one of their lexical properties", where "rules were not much use" (Peter Roach, current editor of Jones' dictionary, personal communication). Later editors have not revised Jones' original stress assignments in any systematic manner (Peter Roach, personal communication).[23] In contrast, Wells (2000) may contain stress patterns that have been assigned in a more systematic manner: John Wells (personal communication) advises that the stress patterns given in his dictionary are not based upon just intuition, but are assigned in a systematic manner based on other phonological evidence.

It is also worth noting that Wells is a much more recently compiled dictionary than Jones: whereas Jones has stress pronunciations dating back to 1913, Wells' pronouncing dictionary was first compiled in 1990. When speaking of her own use of Webster's 1913 English Dictionary, Hay (2003) suggests the general advantage of using an older dictionary when trying to obtain speaker intuitions:

> [T]here may be some serendipitous advantage to using a dictionary which was compiled in 1913. The entries are likely to represent a closer reflection of the dictionary writer's mental representations than more modern-day dictionaries which may be subject to stricter conventions and consistency of content (Hay, 2003: 53).

In sum, it may well be that a more systematic, less intuitive process of stress-assignment is in evidence in Wells (2000) than in Jones (2003); this difference in stress-assignment methods may account for the considerably greater proportion of relative prominence preservation failure in Jones than in Wells.[24] Most importantly, it is desirable that we get successful frequency results with the dictionary we know to most intuitively grounded – Jones (2003) – and we do.

---

[23] A.C. Gimson, who took over from Jones, changed mostly only segmental transcriptions, and subsequent editors (15th edition onwards) have revised stress patterns only on an "ad hoc" basis for individual words ("usually on the basis of their [the editors] intuitions, or in response to advice from someone else") (Peter Roach, personal communication).

[24] Differences between Wells (2000) and Jones (2003) in terms of segmental phonology may not be so clear: Peter Roach (personal communication) advises that, as compared to stress, segmental transcriptions have been more systematically and significantly revised in later editions of Jones' dictionary.

6.5.3 Inconsistent preservation failure

The great majority of words in the non-preserving category do not have just a non-preserving stress-pattern with consistent initial-syllable secondary stress, e.g. ***àn****ticipátion* (*antícipate*) – they also have a preserving pattern, e.g. *antìcipátion* (*antícipate*).[25] This variable behaviour follows from the assumption that the parsing line is not an "absolute" (Hay & Baayen, 2002: 17). Hay (2003) and Hay & Baayen (2002) point out that both whole-word and decomposed access may play a role in the retrieval of words, regardless of which side of the parsing line they fall. It is simply that the further from the parsing line a word gets, the more dominant the direct route or the decomposed route becomes:

> [I]t is not the case that all words falling above the line are definitely and fully decomposed, and all words falling below the line are definitely and completely accessed via their derived forms. On the contrary, we regard decomposition as a continuum (see Hay, 2000 [2003]), and assume that both parsing and whole-word access are likely to play some role in the access of most affixed words (Hay & Baayen, 2002: 217).

Because the parsing line is not an absolute, it is not the case that a word on one side of the parsing line or the other is guaranteed to have either preserving or non-preserving stress: crossing the parsing line simply makes one possibility more likely than the other. Any given word may be retrieved by both the whole-word route, which makes it more prone to non-preserving stress, and the decomposed route, which makes it more prone to preserving stress. Consequently, a word may be recorded with both a preserving and non-preserving stress pattern. Additionally, it is important to remember that a word's place in the decomposition continuum is not the only factor which determines the success of stress preservation – the strength of the default stress pattern plays a part too (§6.5.1, above, and §9.3).

[25] The argument in this section holds whether initial-stressed pronunciations are an outright failure to preserve stress in any form on the second syllable, or just a failure to preserve relative prominence, second-syllable foot-head preservation still being successful: both situations are instances of less than optimal similarity with the stress contour of the embedded word. The possibility of successful foot-head preservation in the absence of successful relative prominence preservation is discussed in §6.5.1 above and in chapter 9.

6.5.4 Other factors which may cause preservation to fail

Since SPE’s proposal for the phonological cycle, stress preservation has been generally assumed to be a “strictly local” phenomenon (Benua, 1997): stress is preserved only from immediately embedded words, e.g. *o**rì**ginálity* (*o**rí**ginal*), and not from more-deeply embedded words, e.g. ****ò**riginálity* (***ó**rigin*). True to this hypothesis of strict locality, more-deeply embedded words have not been considered in the frequency analyses so far in this chapter. In this section, it is argued that stress preservation from more-deeply embedded words should be considered.

6.5.4.1 More-deeply embedded words

Hay & Baayen (2002) point out that relative frequency effects will potentially be complicated when the base of affixation is itself morphologically complex:

> [U]nderstanding the behaviour of *carelessness* may require coming to an understanding of the role of how the token frequency of *careless* relates to the frequency of *careless*, how the frequency of *careless* relates to the frequency of *care*, whether the relationship between the frequency of *carelessness* and *care* is relevant, and whether the relationship between each of these relationships plays any role (Hay & Baayen, 2002: e.n. 2).

From the perspective of stress preservation, it is possible that, in certain frequency contexts, more-deeply embedded words may influence the stress preservation behaviour of the embedding word in question. For example, consistently preserving relative prominence in *auth**èn**ticátion* could be as much due to the influence of *auth**én**tic* as that of *auth**én**ticate*, particularly if *authéntic* is more frequent than *authénticate*.

An area where the effects of the frequencies of more-deeply embedded words are particularly pertinent are the examples which are given in §5.5.2. These examples were words where an apparently non-preserving stress pattern corresponds to the stress pattern of a more-deeply embedded word: e.g. variable secondary stress in ***tò**talitárian~to**tà**litárian* may be the result of preservation (variably) from immediately embedded *to**tá**lity* and more deeply embedded ***tó**tal*. Given that we have already noted in this chapter that stress is more likely to be preserved from more frequent embedded words, it seems reasonable to hypothesise that stress may be preserved from a more-deeply embedded word like *tótal* if it is more frequent than the immediately embedded word. As a test of this hypothesis, the words from Jones (2003) that were given in §5.5.2 are repeated in table 7, along with raw CELEX frequencies (per 17.9 million) of the embedding and putatively embedded words:[26]

	Embedding word (variable secondary stress)	Freq.	More-deeply embedded Word	Freq.	Immediately embedded word	Freq.
1	ambassadorial	4	**ám**bassy (OED)	0	amb**á**ssador	279
2	antipathetic	5	**ànti**páthic (OED)	0	ant**í**pathy	23
3	coincidental	16	c**òin**cíde	189	co**ín**cidence	243
4	humanitarian	38 (adj) 7 (n)	h**ú**man	5113	hum**á**nity	334
5	iconoclastic	12	**í**con	62	ic**ó**noclast	5
6	Mephistophelean	2	Meph**í**sto	0	M**èphi**stópheles	3
7	Shakesperiana	0	Sh**áke**speare	357	Shakesp**é**rian	48
8	totalitarian	98	t**ó**tal	1997	tot**á**lity	58
9	triangularity	0	tr**í**angle	219	tri**án**gular	61
10	triangulation	0	tr**í**angle	219	tri**án**gulate *tri**án**gular?*	0 *61*
11	utilitarian	62 (adj) 2 (n)	**ú**tilise	131	ut**í**lity	154

Table 7: frequency and more-deeply embedded words

For 6 out of the 11 words given in table 7 (*humanitarian*, *iconoclastic*, *Shakesperiana*, *totalitarian*, *triangularity*, *triangulation*), a more-deeply embedded word is of a higher frequency than the immediately embedded word. (For *totalitarian* and *iconoclastic*, the embedding word is also more frequent than the immediately embedded word, independently making preservation from the immediately embedded word less likely.) For an additional two words – *coincidental* and *utilitarian* – the frequencies of the immediately embedded and more-deeply embedded words are quite similar, particularly in the log form argued to be more

[26] *Cèrtificátion~certìficátion* that was given in §5.5.2 has been removed: the two stress variants are argued to have different semantics depending upon whether they are derived from *cértify* or *certíficate* (Jones, 2003).

relevant to lexical access by Hay & Baayen (2002): 5.2 for *coincide* versus 5.5 for *coincidence*; 4.9 for *utilise* versus 5.0 for *utility*. In sum, in light of the frequency information for 8/11 embedding words in table 6, it is plausible that the stress of more-deeply embedded words is affecting the stress pattern of the embedding word.

6.5.4.2 Prefixation

The majority of heavy-initial words which are candidates for relative prominence preservation contain a prefix. Two ways in which prefixation may interact with relative prominence preservation are now explored: (i) the triggering of connected speech stress-shift; (ii) the removal of the environment where preservation failure is possible.

Kenyon & Knott (1944/1953) argue that failure of relative prominence preservation may be due, in part, to connected speech phenomena (repeated from §3.2.1):

> In actual speech, such alternative accentuations such as […] *,impenetra'bility* or *im,penetra'bility*, *in,feri'ority* or *,inferi'ority* are very common, and do not represent more or less desirable pronunciations, but chiefly show **the effect of varying sense stress, emphasis, speech rhythm, semantic distinctions, and other constantly varying factors of connected speech,** so that in many such instances the question which accentuation is preferable is irrelevant (Kenyon & Knott, 1944/1953: xxv [Kager, 1989: 170; see also Halle & Vergnaud, 1987a: 243-4]; boldface SC).

Although we are not working with Kenyon & Knott's pronouncing dictionary, and dictionary entries are supposed to be citation forms, we do have to consider the undesirable possibility that, for some of our dictionary entries, a preserving citation form has not been isolated from connected speech pronunciations by the dictionary editors – if Kenyon and Knott have encountered this problem, then Jones and Wells may have also.

For the sample of words from Jones (2003) which have been extensively examined for frequency effects in this chapter, one factor which may lead to a leftward shift of stress in connected speech is emphasis of a word's prefixation. Bolinger (1961), and subsequently Hay (2003), propose that speakers may use prosody to emphasise a prefix during connected speech if the prefix is clearly perceptible within the word. Hay proposes that pitch accent may be shifted to a prefix for the purpose of contrast:

> Sarah though her cousin was liberal, but I found him completely **il**liberal (Hay, 2003: 91, boldface SC).

Similarly, Bolinger proposes that there may be instances of emphatic prefix stress:

> [Y]ou may **dé**tain them, but don't **ré**tain them (Bolinger, 1961: 89; boldface SC).

In our data, it is possible that emphatic prefix stress overrides relative prominence preservation: e.g. the potential preserving stress pattern of *des**è**gregátion* (*des**é**gregate*) may be overridden by emphatic stress-shift in connected speech:

> We do not need further segregation of boys and girls; we need d**è**segregátion.

In sum, although it is hoped that the dictionary data from Jones (2003) and Wells (2000) is free from connected speech phenomena, there is at least one connected speech phenomenon which could plausibly interfere with relative prominence preservation.

Prefixation may also reinforce relative prominence preservation, by removing the environment where preservation failure is possible. Some of the data for relative prominence preservation contains semantically opaque Latinate prefixations: e.g. *con-* (*conglomeration*). These types of prefixations have a tendency to reject stress (SPE: 94; Liberman & Prince, 1977: 284; Halle & Vergnaud, 1987a: 239; Pater, 2000: 263; Hammond, 2003b), shown by their sometimes reduced vowels: e.g. ***c*[ɒ/ə]*n***_glòmerátion_ (Jones, 2003, s.v. 'conglomeration'). An interaction with relative prominence preservation is consequently expected: if a prefix is independently desired to be unstressed, then it cannot bear initial-syllable secondary stress, and so preservation cannot fail – ****c*[ə̀]*n***_glomerátion_ (*congl**ó**merate*).

Pater (2000) indicates that not all etymologically Latinate prefixations reject stress – semantically transparent ones do not:

> [M]ore semantically transparent cases of prefixation, especially with the very productive prefixes /pre-/, /re-/, /pro-/ and /de-/, do not involve reduction (e.g. *recover* 'cover again' *vs*. *recover* 'get back […]) (Pater, 2000: 263).[27]

There are examples of semantically transparent prefixation with historically Latinate prefixes in the data for relative prominence preservation: e.g. *depopulation* (*de+population*$_N$). Therefore, with semantically transparent Latinate prefixation, there may be no such bar to preservation failure: ***d*[ì:]***populátion* (*depópulate*).[28]

This hypothesis about the effects of semantically opaque versus transparent prefixation is supported by the data for relative prominence preservation in *-ion* words from Jones (2003) (appendix L). There were a total of 70 words which contained etymologically Latinate prefixes (Marchand (1969) was used to check prefix etymology).[29] Latinate prefixations of bound bases were categorised as semantically opaque; Latinate prefixations of free words were categorised as semantically transparent.[30] In a chi-square analysis, the non-preserving category is characterised by containing significantly fewer prefixations of bound bases than the preserving category – $\chi^2 = 37.379$, 1df, $p<.001$. The chi-square contingency table is given in (3):

[27] *Pace* Pater (2000), I do not hold that the simple absence of vowel reduction definitely indicates the presence of stress (chapter 2). What is important is that for vowel reduction to occur in the prefix, we know the prefix must absolutely be de-stressed.

[28] Fidelholtz (1975) also points out that light initial syllables of words quite freely reduce: perhaps this reinforces second-syllable preservation in light-initial words – schwa is unstressable – in part contributing to the more consistent second-syllable preservation in this group.

[29] The prefixes *em-* (*en-/im-/in-*) (e.g. *incineration*, *encapsulation*) have been classed as Latinate following Halle & Vergnaud (1987a) and Pater (2000). Marchand (1969: 163-4) does point out that, in some words with *em-*, it is difficult to tell whether the original source was Latinate or native, but that in "learned words the prefix can safely be considered Latin". 'Learned' seems an appropriate description of the preservation examples being analysed here.

[30] This is a very broad division between the group where vowel reduction may occur – opaque Latinate prefixations – and the group where it is systematically absent. Pater (2000) shows that there is a great deal of variation amongst the opaque Latinate prefixations as to whether or not vowel-reduction occurs in the prefix.

An example is counted as being a 'prefixation of a free word' only if the rest of the embedding word is a free word that is not marked as rare or obsolete in the OED online, e.g. *de+population*$_N$, and there was a very transparent semantic relationship between the base of prefixation and the prefixed word.

(3) Relationship between prefixation and preservation behaviour

		Stress	
		Preserving	Non-preserving
Latinate prefixation	Free-word base	1 (11.0) 2.1%	15 (5.0) 68.2%
	Bound base	47 (37.0) 97.9%	7 (17.0) 31.8%

$\chi^2 = 37.379$, 1df, $p<.001$

The theoretical machinery for handling the variable metrification behaviour of Latinate prefixations is presented in §9.3.2.1.

6.6 Conclusion

This chapter has presented the most interesting empirical finding of the thesis: the observation that relative prominence preservation is probabilistically dependent upon word frequency. This is a first for left-edge stress preservation, and the first examination of Hay's (2001, 2003) hypothesis of relative frequency with respect to stress preservation in general. Effects of absolute frequency have been found: relative prominence preservation is more likely to fail if the embedding word is more frequent, and/or the embedded word is less frequent. We have also discovered word-specific relative frequency effects: relative prominence preservation is more likely to fail when the embedding word is more frequent than the embedded word, but less likely to fail if the embedding word is less frequent than the embedded word.

With this chapter, the empirical investigation into weak stress preservation concludes. Chapters 4 and 5 have shown that weak stress preservation is a reality with which a model of English phonology must deal; the problems this causes for monostratal Optimality Theory, and the subsequent defence of a model of Stratal Optimality Theory, are presented next in chapter 7. The present chapter has told us that any weak preservation mechanism must be probabilistic. The means by which this is achieved in Stratal OT is the mechanism of fake cyclicity; this proposal is presented in chapter 8.

Part III: theoretical analysis

Chapter seven: Stratal Optimality Theory for English stress preservation

7.0 Introduction

The existence and behaviour of English weak stress preservation has been established in the preceding chapters. In this chapter, it is shown that Stratal Optimality Theory ('Stratal OT') is the means by which English stress preservation should be handled in OT.

The argument for handling English stress preservation in Stratal OT necessarily consists of two parts: first, the defence of OT in general for handling stress preservation; second, the defence of Stratal OT in particular. When it comes to defending specifically Stratal OT, I draw not only upon the advantages of Stratal OT over other models of OT with particular respect to the handling of stress preservation, but also upon some of the advantages of Stratal OT over other models of OT in the wider context of phonological opacity.

This chapter is structured as follows. In §7.1, the advantages of handling stress, and stress preservation in particular, in OT are shown. In §7.2, the long-recognised problem of phonological opacity in OT is presented, and the relevance of this issue to English stress preservation is demonstrated. The ways in which the type of phonological opacity characterised by English stress preservation can be handled in OT are introduced: Output-Output Correspondence Theory, which has received the bulk of the attention in OT to date, and Stratal OT. In §7.3, Output-Output Correspondence Theory is discussed in detail, including a particular version which has been proposed, in part, to deal with English stress preservation: the Transderivational Correspondence Theory of Benua (1997). The alternative solution offered by Stratal OT is then introduced in §7.4, before the chapter concludes in §7.5.

7.1 Why OT is good for stress

An introduction to OT was given in §1.3. There, we saw that OT differs from rule-based theories of Generative Phonology by virtue of its output-orientation: OT phonology consists entirely of violable constraints which make statements about the well-formedness of outputs. In contrast, rewrite rules operate in an 'output-blind' fashion: rules apply without respect as to whether they are creating a well-formed output or not. OT's output-orientation makes it preferable to rules when it comes to the general handling of stress.

The problem with the operation of rules for stress is that their output-blindness means that they incorrectly or over-generate metrical structure; this then has to be repaired by rules such as various types of destressing (for examples see: SPE; Liberman and Prince, 1977; Kiparsky, 1979; Hayes, 1981) and stray syllable adjunction (Kager, 1995b: 20) in order to achieve the correct surface form. In contrast, in OT, no repairs are ever necessary: the incorrect structure is not generated in the first place (Benua, 1997: 11). Rather, the well-formed output is targeted by the constraints from the outset (Kisseberth (1970) and 'rule conspiracies').[1]

Even within primarily rule-based models, it was recognised that constraints upon outputs were required to deal with metrical phenomena, e.g. Liberman's (1975) 'no-clash' (Kager, 1999: 56-7), Selkirk's (1984: 109) 'Anti-Lapse Filter', and Halle & Vergnaud's (1987a: 238) 'Stress Well' – *CLASH-HEAD in OT. Indeed, as the following quotation from Alber (1998) indicates, surface-wellformedness constraints seem to be inevitable in metrical theory:

> If we want to give the derivational analysis the same restrictiveness that the constraint based analysis has, we would have to specify that destressing rules can resolve clash only in such a way that the final output does not disturb 'important' requirements on the stress pattern of the language (e.g. for German: initial syllables must be stressed). But this would just mean to introduce wellformedness constraints on the output into the rule system (Alber, 1998: 130).

Clearly, output-wellformedness constraints are necessary in any metrical theory. OT's advantage over any sort of 'mixed' model which uses rules as well as constraints is that OT's theoretical machinery is more simple and economical: OT uses only constraints (Kager, 1999: 57).

With respect to metrical phenomena, the parallelism of OT's constraint evaluation, as well as its surface-orientation, is advantageous. A well-known

[1] As noted by Booij (1996: 69), output-orientated constraints are not OT's unique innovation, but are also seen in the Obligatory Contour Principle of Goldsmith's (1976) Autosegmental Phonology, in Harmonic Phonology (Goldsmith, 1990, 1993), and Declarative Phonology (Coleman, 1995 a, b; Scobbie et al. 1995).

criticism of rule-based theories is that the ordering of certain rules is just stipulated to achieve the correct eventual result (e.g. the ordering of stress assignment before epenthesis in Levantine Arabic (Kiparsky (2000: 353), citing original work by Brame (1974)). Rule ordering runs into significant problems when it seems that two processes must necessarily be mutually unordered. In contrast, OT's parallel constraint evaluation can naturally cope with mutually unordered interactions. Consider, for example, the crucially parallel interaction of phonological processes encountered in Pater's (1995: 9-10) analysis of the interaction of sonorant-coda coalescence with stress in OT. This analysis indicates that stress and syllabification must crucially occur in parallel. Pater shows that whether the second syllable of an Hσ pre-tonic sequence is stressed depends upon whether the final consonant of the second syllable can function as nucleus: *Pènns***[l̩]***vánia* (syllabic sonorant in an unstressed syllable) versus *Tìmb***[ʌ̀k]***tóo* (obstruent coda of a stressed syllable) (see §2.3.2). However, whether or not a sonorant is syllabic or functions as a coda depends upon whether the syllable is stressed: *Pènns**l**vánia* versus *b**àn**dána*. This circular reasoning would clearly be difficult to capture in a rule-based theory, where syllabification and stress assignment must be ordered with respect to one another. In contrast, in OT, it is anticipated under the parallel interaction of constraints, as shown in (1) (based on Pater, 1995: 9-10):

(1) Parallel interaction of syllabification and stress-assignment in OT

WEIGHT-TO-STRESS = heavy syllables are stressed

*SONNUC = no sonorant nuclei

*OBSNUC = no obstruent nuclei

STRESSWELL = no stressed immediately pre-tonic syllables

PARSE-σ = all syllables belong to a foot

Inputs: /*Pennsylvania*/, /*Timbuctoo*/, /*bandana*/	PARSE -σ	*OBS NUC	WEIGHT -TO-STRESS	STRESS WELL	*SON NUC
☞ a. (Pènn.sl) │ vánia					*
b. (Pènn)(sỳl) │ vánia				*!	
a. (Tìm.buc) │ tóo			*!		
☞ b. (Tìm)(bùc) │ tóo				*	
c. (Tìmbc) │ tóo		*!			
☞ a. (bàn) │ dána				*	
b. bn │ dána	*!				*
c. ban │ dána	*!		*		

In conclusion, OT's output orientation offers a more constrained, economical and undoubtedly simpler analysis of stress compared with previous rule-based approaches. OT's parallelism also has the potential to avoid rule ordering paradoxes. It will now be shown that, in some respects, OT has already shown itself to be particularly desirable for the analysis of stress preservation.

7.1.1 Stress preservation

By virtue of its theory of constraint interaction, OT is able to offer a highly restrictive analysis of stress preservation. The OT analysis recognises preservation as a violable constraint which interacts directly with markedness constraints. In doing so, the OT analysis can capture the fact that words which are candidates for preservation are subject to exactly the same markedness conditions as words which are not candidates. This not only offers a restrictive theory of phonology, but allows the straightforward prediction of phonological environments where preservation will manifest, e.g. *orí̲ginal* → *orì̲ginálity*, *hó̲pe* → *hó̲pelessness*, and those where it will not, e.g. *á̲tom* → *a̲tómic*.

OT's handling of stress preservation as a violable constraint is prefigured in Burzio's (1994) constraint-based, non-OT analysis of English stress. Following Selkirk (1980), Burzio recognises that stress preservation will not occur across the board, but only where it does not threaten the satisfaction of a higher priority, purely phonological requirement (Kager, 1995b). This concept is captured in Burzio's principle of 'Metrical Consistency', which takes the form of a violable constraint:

(2) Metrical Consistency (Burzio, 1994: 228)

Every morpheme must be as metrically consistent as possible.

Burzio's recognition of stress preservation as a ranked and, vitally, **violable** requirement is an important step forward in the modelling of stress preservation: if preservation is violable, no repairs are required to correct preservation which results in ill-formed feet, such as Hammond's (1989: 147-8) proposed destressing of the initial syllables of the embedding forms in **pàréntal* (*párent*) and **mòdérnity* (*módern*). Additionally, by showing that Metrical Consistency interacts directly with phonological well-formedness constraints that are applicable to all words, simplex and complex alike, Burzio's theory can also capture the empirical situation whereby "morphologically-simple and morphologically-complex words differ only minimally and yet crucially in their stress patterns" (Burzio, 1994: 12).

Burzio's insight is retained in the OT handling of stress preservation. In OT, faithfulness constraints are the incarnation of Burzio's theory of Metrical Consistency. Like Metrical Consistency, faithfulness constraints are violable, and they interact directly with the markedness constraints that state purely phonological requirements. The same ranking of markedness constraints will be applicable to complex and simplex words alike, so that, as in Burzio's theory, constraint evaluation "puts a limit on how deviant the deviant phonology of complex words can be" (Benua, 1997: 6; see also Alber, 1998: 131). An example of this is given for English weak stress preservation in the tableau in (3); (3) is based upon the constraint ranking established for English by Pater (2000):[2]

(3) Evaluation of morphologically simple and complex words with the same constraint hierarchy

FTBIN = feet are binary at moraic or syllabic level

PARSE-σ = syllables belong to a foot

IDENT-STRESS = preserve any stress present in the input; score one violation for demotion of stress from primary to secondary; score two violations for failure to preserve stress at any level

[2] The analyses in both (3) and (4) are simplified for the present purposes.

ALIGN-L = align the left edge of the prosodic word with the left edge of a foot

Inputs: /*Tatamagouchi*/, /*oríginal-ity*/	FTBIN	PARSE-σ	IDENT-STRESS	ALIGN-L
☞ a. (Ta.ta)ma │gouchi		*		
b. Ta(ta.ma) │gouchi		*		*!
c. (Ta.ta.ma) │gouchi	*!			
a. (o.ri)gi │nality		*	**!	
☞ b. o(ri.gi) │nality		*	*	*
c. (o.ri.gi) │nality	*!		**	

Both the simplex word *Tàtamagóuchi* and preserving *orìginálity* are the products of exactly the same constraint ranking. The only difference is that the constraint which ensures stress is preserved, IDENT-STRESS, is vacuously fulfilled in *Tàtamagóuchi* because there is no stress to preserve, whereas, in *orìginálity*, fulfilment of IDENT-STRESS violates ALIGN-L.

Because preservation is a violable constraint which interacts with markedness constraints, we can directly predict the contexts in which stress will not be preserved, e.g. *átom* → *a̲tómic*. Pater (2000) argues that the constraint ALIGN-HEAD (McCarthy & Prince, 1993b), which requires primary stress to be as close to the right edge of the word as possible, is ranked between FT-BIN and PARSE-σ from (3). This new ranking (FT-BIN>> ALIGN-HEAD>> PARSE-σ>>IDENT-STRESS>>ALIGN-L) predicts the stress pattern of a monomorphemic word like *banána*, but can also account for non-preservation from *átom* in *a̲tómic*. (Not shown in (3) is the constraint NONFIN; this constraint renders the final syllable of *banana* extrametrical by dominating PARSE-σ. In the case of *-ic* adjectives like *atomic*, the final syllable is footed; I follow Bermúdez-Otero & McMahon (2006) in arguing that, here, the ranking of NONFIN and PARSE-σ is reversed, something possible under Bermúdez-Otero & McMahon's cophonology approach – see §9.3.2. Following Bermúdez-Otero & McMahon, I assume that final-consonant extrametricality still applies in *atomic*. Finally, I assume that a constraint which requires feet to be left-headed, TROCH, is undominated in English, ruling out **(ba.ná)na*.)

(4) Markedness overrules stress preservation

Inputs: /*banana*/, /*átom-ic*/	FTBIN	ALIGN-HEAD	PARSE-σ	IDENT-STRESS	ALIGN-L
☞ a. ba(ná.na)		*	*		*
b. (bá.na)na		**!	*		
c. (bà)(ná)na	*!	*			
☞ a. a(to.mi)c		*	*	**	*
b. (á.to)mic		**!	*		
c. (à)(tó.mi)c	*!	*		*	

The high ranking of FTBIN in (4) rules out degenerate feet consisting of a single light syllable – *(Ĺ). Consequently, preservation from *átom* to give **àtómic* is ruled out: taking into account the requirements of both FTBIN and ALIGN-HEAD, preservation would necessitate the creation of a degenerate foot on the preserving initial syllable. The combined effects of FTBIN and ALIGN-HEAD similarly rule out the ill-formed monomorphemic candidate c. **bànána*.[3]

In sum, by recognising stress preservation as a violable constraint which interacts directly with markedness constraints, OT is able to offer a restrictive analysis of stress preservation. Complex and simplex words are subject to the same ranking of constraints, and the only potential for the phonology of complex words to differ from that of monomorphemic words is through the interaction of the preservation constraint with markedness constraints. OT can also offer a simple, principled account of why stress preservation fails in some phonological contexts, but not others.

7.2 OT's opacity problem

We have seen that OT's ability to formalise stress preservation as a violable constraint enables a restrictive analysis of stress preservation. Yet, perversely, while OT's parallel constraint evaluation permits such elegant analysis, it is the strong assumption of parallelism by many OT practitioners that causes problems for OT

[3] With just a single stress identity constraint, this analysis therefore makes the important prediction that exceptional stress will not occur in monomorphemic words in the same environments as preserving stress will not occur in complex words. This issue is important in the defence of Stratal OT, and is discussed further in §7.4.1.1, below.

with respect to phonological opacity (for a good critique, see Idsardi (2000)). This is important, as English stress preservation is an instance of phonological opacity.

The term 'phonological opacity' describes a situation whereby, in a surface form, a phonological generalisation does not apply in the environment where it is predicted to – phonology 'misapplies'. However, in cases of opacity, this misapplication of phonology in the surface form is not simply an aberration, but is systematic and predictable: the phonological generalisation which is opaque in a surface form becomes transparent if we look at a level of representation below the surface (Kager, 1999: 372).[4] Opaque generalisations take two forms: 'overapplication', where a process applies even though it is not conditioned by the surface environment, and 'underapplication', where a phonological process does not apply where it is conditioned (Wilbur, 1973). Non-opaque phonology is produced by 'normal application' of phonology in the surface form.

Following Bermúdez-Otero (in preparation), two types of phonological opacity are recognised over the following chapters: that which is confined to a particular morphological or syntactic context – 'paradigmatic opacity' – and that which occurs in all realisations of a morpheme, regardless of its morphosyntactic context – 'non-paradigmatic opacity'.[5] A discussion of non-paradigmatic opacity will be left until later in the chapter (§7.4.2), where it is relevant to the defence of Stratal OT. For now, just paradigmatic opacity will be considered, as stress preservation can be understood as an instance of this type of opacity. Take, for example, the preserved second-syllable secondary stress in *orìginálity*. This stress pattern is not what our surface phonological generalisations predict for this phonological environment: the expected initial dactyl, as in *àbracadábra*, underapplies (Benua, 1997: §5.2). But the stress pattern of *orìginálity* ceases to be opaque when we take into account its

[4] For the original source of the use of 'opacity' in this phonological sense see Kiparsky (1971, 1973: 79). Bermúdez-Otero (in preparation) gives a good recent overview of phonological opacity.
[5] As is evident from the definition of 'paradigmatic opacity', Bermúdez-Otero's use of the term 'paradigmatic' is quite loose – his usage does not imply that paradigms have any particular status in morphological theory (Ricardo Bermúdez-Otero, personal communication), in contrast with Output-Output Correspondence Theory. In most recent OT literature (e.g. Benua, 1997; Kiparsky, 2000; Burzio, 2005b), a distinction is explicitly made between terms such as 'paradigmatic identity' or 'misapplication', and 'opacity', where the former is equivalent to our 'paradigmatic opacity', and the latter to our 'non-paradigmatic opacity'. Bermúdez-Otero's terminology is more representative, in that it captures the fact that both phenomena are instances of misapplication conditioned by intermediate representations; the only difference is in what type of representations these are, as discussed in §7.4.2 below.

morphological structure: *orìginálity* contains the embedded word *oríginal*, and in *oríginal*, primary stress has applied normally – its application is transparent. Normal application at this intermediate level of representation, and the morphological structure of *originality*, predicts opaque stress as in *orìginálity*. Stress preservation is therefore an instance of paradigmatic opacity.

We will now see that OT's strong emphasis upon parallelism makes the theory too restrictive to be able to cope with phonological opacity, and therefore stress preservation. Part of classical OT's strength is that it permits only a single, direct mapping from input to output. (The term 'classical OT' is used to refer to a monostratal model of OT which does not include any correspondence constraints except for Input-Output Correspondence, outlined in §1.3.) The form of this mapping is shown in figure 1:

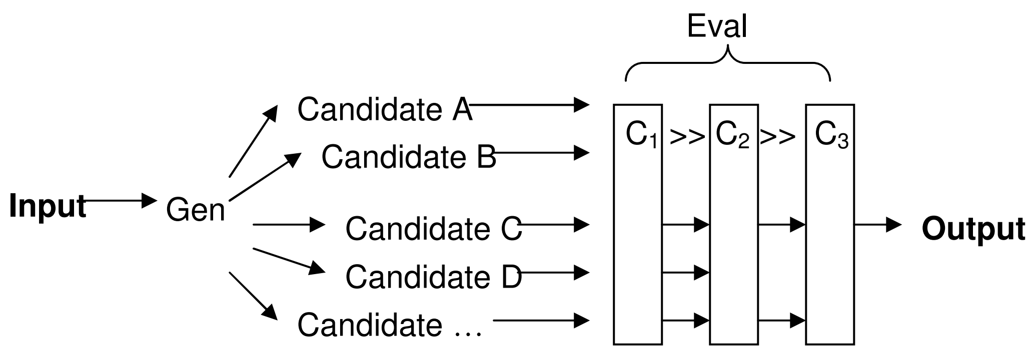


Figure 1: classical OT grammar (based on Kager, 1999: 9)

Classical OT therefore avoids the problems of a theory like SPE, where a "potentially unlimited number of intermediate stages must be recognized" in a phonological derivation (Kenstowicz, 1996: 369), resulting in independently unmotivated intermediate representations. In classical OT, there are just two levels of representation: Input and Output. However, classical OT's restrictiveness means that it suffers a serious setback when confronted by a situation in which intermediate representations between input and output seem to be genuinely necessary: the paradigmatic opacity of English stress preservation.

Classical OT absolutely requires the rejection of the phonological cycle: the cycle is a serial mechanism which formulates intermediate representations. The result of this restriction is that simplex and complex words will be incorrectly predicted to have exactly the same stress patterns (see also §1.3.2). For example, preserving stress in *orìginálity* crucially relies upon the stress of *oríginal* having been assigned in advance. Two passes through the phonology are therefore required, as argued by the cycle: the first to assign stress to give *oríginal*, and the second which builds upon this word to give *orìginálity*. However, if the assumed strong parallelism of classical OT is to be respected, the step-by-step derivation proposed by the phonological cycle simply cannot be incorporated into the theory. Under the single pass through Gen and Eval required by strong parallelism, the stress pattern of *original* cannot be assigned in advance of that of *originality* – both must be derived simultaneously. Consequently, the incorrect output **òriginálity*, paralleling monomorphemic *àbracadábra*, is predicted: there is no intermediate representation *oríginal* which can cause the initial dactyl to underapply to give preserving *orìginálity*.

One way around classical OT's opacity problem, which would not require the introduction of intermediate representations (nor any amendments to the OT constraint inventory) would be to propose that stress is marked on the input of complex words like *originality*, i.e. /*o(rígi)nality*/. With the right constraint ranking, these underlying foot-heads would be preserved in the output by the action of Input-Output faithfulness, and *orìginálity* would be achieved without introducing any levels of representation other than input and output: the stress of *oríginal* would not need to be assigned in advance. However, this analysis appears to miss the vital generalisation that the stress pattern of *orìginálity* relies upon that of *oríginal*. It was shown in chapters 4 and 5 of this thesis that there is very strong empirical evidence for English stress preservation: we cannot simply write off the misapplication of secondary stress in complex words as idiosyncratic distinctive stress present in a word's underlying representation.[6]

In sum, English stress preservation is a major problem for classical OT. Tellingly, Hammond (1999a) largely avoids the treatment of complex words in his book-length OT work on English prosody, focusing primarily upon the 'distributional regularities' of monomorphemic words (Honeybone, 2000: 166). In the seven pages where Hammond deals with stress preservation (1999a: 322-329), he either sacrifices the preservation generalisation by resorting to lexical diacritic

[6] Fake cyclicity, discussed in chapter 8, proposes underlying stress of sorts, but does not sacrifice the weak preservation generalisation.

accents instead of a preservation mechanism, or else is forced to make serious modifications to the theory of OT, such as Smolensky's (1993) Local Conjunction.[7]

There are two main routes by which OT has attempted to solve its opacity problem which are relevant to the paradigmatic opacity of stress preservation:[8]

1. Enforcing paradigmatic opacity without introducing an intermediate level of representation between input and output, thus retaining classical OT's strong parallelism.
2. Permitting an intermediate level of representation to enforce paradigmatic opacity, i.e. the phonological cycle, thus proposing a serial analysis at odds with classical OT's strong parallelism.

Route 1 – finding a means of handling opacity which retains classical OT's strong emphasis on parallelism – has received by far the bulk of attention in OT to date, seemingly on the assumption that strong parallelism is a defining and desirable characteristic of OT.[9] Nevertheless, route 2 has also been attempted: serialism, as evident in cyclic phonology (SPE; Halle and Vergnaud, 1987a) and Lexical Phonology (e.g. Kiparsky, 1982; Giegerich, 1999), has also been introduced into OT. This latter route is clearly in conflict with classical OT's assumption of strong parallelism, and is the route represented by Stratal OT and defended later in this chapter. However, before we can consider Stratal OT, route 1 – the attempt to retain classical OT's strong parallelism in the face of phonological opacity – must be considered and shown to be unsatisfactory.

[7] Local Conjunction disrespects the strict domination of constraints.

[8] Generalised Alignment Theory (McCarthy & Prince, 1993b) is also a means of representing morphologically conditioned stress (e.g. Kenstowicz, 1995) without resorting to intermediate representations. However, there are instances that this theory cannot handle where the cycle seems to be genuinely required (Kenstowicz, 1995: §4, 1996: 366-7). To my knowledge, English stress preservation is one such case.

[9] Both Bermúdez-Otero (1999, in preparation) and Orgun (1994, 1996a) suggest that the high premium placed upon strong parallelism by many OT practitioners is due to the now widely accepted belief that parallel processing is psychologically more plausible than serial processing. However, precisely what OT practitioners mean by 'parallelism' needs serious consideration, as shown by the discussion of Benua (1997) in §7.3.1.3 of this chapter.

7.3 Paradigmatic opacity, route 1: Output-Output Correspondence Theory

Output-Output Correspondence Theory has been proposed in order to model paradigmatic opacity in OT while simultaneously respecting strong parallelism.

In chapter 1, McCarthy & Prince's (1995, 1999) Correspondence Theory of phonological faithfulness was briefly outlined. This theory proposes that faithfulness constraints play a part in an optimality-theoretic grammar, either between input and output, or different representations co-present in the same output (reduplication), or, vitally (in developments upon McCarthy & Prince's work), between separate outputs: Output-Output Correspondence Theory.

Output-Output Correspondence Theory ('OO-correspondence') has been used to model supposed paradigm uniformity, including truncation (Benua, 1995), phonological relationships between affixed words and their non-affixed base (Burzio, 1996; Kenstowicz, 1996), as in stress preservation, and paradigm levelling (Kenstowicz, 1996).[10] (OO-correspondence constraints also go by the name of 'Anti-Allomorphy' constraints in Burzio (1996), and simply 'identity' constraints in Kenstowicz (1996).)

OO-correspondence meets classical OT's requirement of strong parallelism. Identity relationships or 'correspondences' between independently occurring outputs apply laterally, as shown in figure 2:

OO = OO-correspondence constraint

IO = IO-correspondence constraint

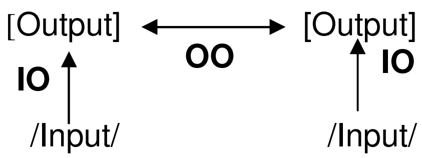


Figure 2: OO-correspondence[11]

[10] The incorporation of paradigm uniformity into formal phonological theory was proposed much earlier than OT, in Harris (1973).

[11] In line with Benua (1997: 27), I assume that all outputs have their own inputs (cp. the diagram shown in Kager, 1999: 275).

No intermediate level of representation between input and output need be introduced to trigger opaque phonology in the output: the representations required to induce opacity are already present as outputs.

Under OO-correspondence, identity relationships can only hold between independent outputs – words. As noted by OO-correspondence's advocates, this offers a greater restriction than the cycle of earlier generative theories, where the cycle was not, by definition, restricted to words (Kager, 1999: 285) (although see §7.4.3.2). For example, Kiparsky (1982) and Mohanan (1986) both propose that there is a pre-morphology cycle in Lexical Phonology, so that possible inputs to the first cycle include bound roots, even though it is widely recognised that bound roots are not phonological domains. Some proponents of OO-correspondence propose that identity relationships may hold between subconstituents of words, e.g. affixes (Kenstowicz, 1996: 382; Burzio, 2000, 2005a, b), although the overall domain of OO-correspondence is still the word.

A key issue with respect to OO-correspondence is that OO-correspondence relationships intrinsically reinforce identity between **any** independently occurring words: unlike the cycle, OO-correspondence relationships are not automatically restricted to reinforcing identity effects between pairs of words where one word is an immediate subconstituent of another. Because they are not connected to any sort of serial derivation, OO-correspondence relationships may hold between an embedding word and its more deeply embedded constituent, as between *origin* and *originality*; or between words only related by sharing the same affix (e.g. *originality*, *subliminality*, *sanity*); or, even, between words which are not morphologically related at all (Kenstowicz, 1995: 433; Orgun, 1996a: §3.5; Benua, 1997: 31-2; Kager, 1999: 287; Bermúdez-Otero, 1999: §3.4.1.3, 2007b; Burzio, 2005b: 67). Clearly, without further restriction, OO-correspondence opens up a can of worms. In contrast, the phonological cycle is highly restrictive, proposing that only forms which are 'strictly local' to one another have direct morphophonological relationships: e.g. *original* and *originality*, but not *origin* and *originality*.

OO-correspondence relationships are also inconclusive as to the direction of morphophonological identity relationships. By virtue of the serial nature of the cyclic derivation, the cycle only predicts a directional identity effect from embedded to embedding word. In contrast, the null hypothesis of OO-correspondence is that identity effects are symmetrical: the stress pattern of *oríginal* could influence that of *origin* (i.e. **orígin*). No good evidence for such backwards application of stress preservation has been reported in the literature to date. Burzio (1994) argues that stress preservation should be a symmetrical relationship and so also from embedding to embedded word, and proposes that this is supported by *ànat**ó**mic* ← *ànat**ó**mical* and *inh**á**bit* ← *inh**á**biting*. However, these examples provide no evidence for back-copying of stress from the embedding to the embedded word: *inhábit*$_v$ is simply the result of normal stress application, as verbs are not subject to final-syllable extrametricality (Hayes, 1982); *ànatómic* is the result of the established non-extrametrical behaviour of -*ic*, cf. *Milt**ó**nic* (← **Milt**ó**nical*).[12] It is therefore unsurprising that Burzio later retracts his claims for back-copying preservation (Benua, 1997: 241). Kraska-Szlenk (2007: 140) also notes the absence of back-copying stress preservation English, pointing out that, even when embedding words are much more frequent than the embedded words, and hence back-copying might be presumed to be likely, it does not occur.[13]

Nevertheless, some OT practitioners have embraced the possibility that OO-correspondence relationships may form a wide-ranging network of phonological relationships, notably Burzio (1996, 2005a, b). Similarly, when modelling the process of paradigm levelling, Kenstowicz (1996) employs a principle of 'Uniform Exponence' to encourage general uniformity among all members of a paradigm:

(5) Uniform Exponence (Kenstowicz, 1996: 382)
[M]inimize the differences in the realization of a lexical item (morpheme, stem, affix, word).

Uniform Exponence permits a symmetrical relationship between the words being evaluated – a derived word can affect the phonology of a base word as well as vice-versa – and neither word need necessarily be a constituent of another.

---

[12] As pointed out by Burzio (1996: 133) himself, "-*ic* is metrically bisyllabic quite generally (yielding presuffixal stress), and not just in the cases in which there is an -*ical* variant".
[13] Further supporting the unidirectionality argument from the perspective of analogy, Kraska-Szlenk (2007: 137) notes that it is "hard to come up with unquestionable evidence" for any cases of bidirectional phonological identity, not just examples related to stress preservation.

In spite of the arguments like Burzio's and Kenstowicz's, the unconstrained nature of OO-correspondence has met with criticism. Problematically for OO-correspondence, it does indeed seem to be the case that the strong arguments for identity between independent words are the asymmetrical, strictly local relationships captured by the traditional cycle. Benua (1997) makes convincing arguments for both truncation and base-affix relationships being best-handled with strictly local, asymmetrical identity relationships. Kenstowicz (1996) argues that a more general principle of Uniform Exponence is required for paradigm levelling, as mentioned, but Hale et al. (1998) show that Kenstowicz's principle of Uniform Exponence is problematic: it is not clear what counts as a 'realisation of a lexical item', making any analysis invoking Uniform Exponence unrestrictive. Kiparsky (1998b) also proposes an alternative analysis where Uniform Exponence is not necessary (see also Orgun, 1996a, b; Kiparsky, 1998a; Bermúdez-Otero, 1999: 118). One instance where there **is** general agreement that symmetrical identity is necessary is reduplication (although see Kiparsky, 2007b). However, as pointed out by Benua (1995, 1997), reduplication is an instance of intra- rather than inter-word identity, and so fall outside the remit of OO-correspondence.[14] In short, the evidence indicates that OO-correspondence requires external restriction.

As we will see in §7.3.1, Benua (1997) tries to constrain OO-correspondence theory by strictly constraining – by stipulation – the contexts to which OO-correspondence relationships are applicable. Benua (1997: 33) argues that the only phonologically relevant paradigms are both "linear" and "local", just as is proposed by the cycle. In restricting OO-correspondence in this way, Benua's theory builds upon the notion of Base-Identity already present in OO-correspondence:

[14] Bermúdez-Otero (1999) argues that Stratal OT could cope with the symmetrical identity of reduplication, cp. Benua (1997: 51): "Since B[ase]R[eduplicant]-identity constraints impose conditions of maximum harmony on the output representation of reduplicated words, they can be classed as markedness constraints [...]; the fact that markedness in this case is defined as disparity between two constituents of the target representation is irrelevant, for the relationship between these constituents is syntagmatic, rather than transderivational" (Bermúdez-Otero, 1999: 121). More recently, there have been serious challenges to theories of reduplication based on Base-Reduplicant Correspondence Theory. Inkelas & Zoll (2005) show that Base-Reduplicant Correspondence predicts unattested types of opaque reduplication phonology and that, vitally, apparent symmetrical identity, with what seems to be the copying of phonological characteristic from reduplicant to base, is not due to identity effects between the base and reduplicant. Kiparsky (2007b) also challenges the argument that reduplication results from Base-Reduplicant Correspondence.

> [A] constraint of Base-Identity evaluates candidate outputs for one word in terms of their similarity to the morphologically related base from which the word is derived […] The Base-Identity constraint [...] is relevant to situations in which one word is an **immediate constituent** of a larger derivative structure (Kenstowicz, 1996: 364, 390; boldface SC).[15]

The vital difference is that, whereas Base-Identity is just one of many different types of OO-correspondence relationships in analyses like Kenstowicz (1996), Benua argues that strictly local Base-Identity is the **only** sort.

If the extra empirical scope offered under the null hypothesis of OO-correspondence is not required, and OO-correspondence is simply stipulated to apply in the same paradigmatic contexts as the phonological cycle, the obvious question is whether Benua's theory really offers an improvement upon the phonological cycle, and whether a cyclic version of OT – Stratal OT – would not be preferable. Benua's proposal of OO-correspondence will only be preferable to Stratal OT if we believe, as Benua must, that the cycle is either fundamentally incompatible with OT, and/or that incorporating the cycle into OT creates disadvantages compared to the introduction of a stipulatively restricted form of OO-correspondence into OT. In §7.3.1 and §7.4, it is shown that neither is the case.

7.3.1 Benua (1997)

Benua's (1997) theory of OO-correspondence – Transderivational Correspondence Theory ('TCT') – will now be outlined with particular reference to its handling of English stress preservation.[16]

Benua (1997) is not the only treatment of stress preservation with OO-correspondence: OO-correspondence analyses are also presented in Kenstowicz (1995), Pater (1995, 2000), Burzio (1996: 136, 1999: 1, 2000, 2005b), Alber (1998) (for stress preservation in German), and Kager (2000). TCT is concentrated on here

---

[15] This form of Base-Identity – whereby faithfulness holds between immediate sub-constituents alone – is not accepted by everyone. For example, Kager (2000) proposes a 'transitive' Base-Identity relationship, i.e.: *origin* is the base of *original*, and *original* is the base of *originality*, and so *origin* is also the base of *originality*. This transitive version would make the wrong predictions under the accepted empirical beliefs about English stress preservation, hence Benua's (1997) more constrained version of Base-Identity. However, see the discussion of stress preservation and the 'Manchester Paradigm' in chapter 8.

[16] Some of the ideas discussed in the following critique of Benua (1997) are presented in Collie (2005).

because Benua gives a comprehensive treatment of English stress preservation, and is explicit about the stipulations that must be imposed upon OO-correspondence if the accepted beliefs about English stress preservation are to be captured. It has already been seen in this thesis that these beliefs are not incontrovertible, but, as this is an issue for a cyclic analysis as well as Benua's theory of TCT, these issues are left for separate treatment in the following chapters.

7.3.1.1 Strict locality and linearity

In TCT, an OO-correspondence relationship mirrors every morphological derivation, e.g. affixation, truncation. OO-correspondence relationships hold between base and derived word. In TCT, as in Halle & Vergnaud (1987a), affixes are diacritically marked for which phonological processes they trigger.

Benua proposes two clusters or 'meta-rankings' of faithfulness constraints: OO_1 and OO_2. The faithfulness constraints that cluster under OO_1 produce the phonological effects associated with affixation with class 1 suffixes: e.g. stress shift and weak stress preservation. The faithfulness constraints that cluster under OO_2 bring about the effects of class 2 affixation: stress neutral suffixation and strong stress preservation.

The clustering of faithfulness constraints into two meta-rankings – analogous to the two strata of Lexical Phonology – does not fall out of the theory of OO-correspondence: unlike the arguments made in stratal models, no morphological properties like affix ordering, category of the affixation base, or productivity determine which phonological class a morphological construction belongs to. The existence of two phonological classes is stipulated entirely upon the observation that affixes appear to trigger two distinct types of phonological behaviour in the bases to which they attach, and an affix arbitrarily belongs to either (Benua, 1997: 166). This is an important issue which is returned to later in this chapter (§7.4.1).

Benua proposes that all affixes are associated with a sub-categorisation frame (Lieber, 1980; cp. Giegerich, 1999).[17] One of the functions of the subcategorisation

[17] Giegerich (1999) proposes that full listing applies to stratum one morphology, and subcategorisation frames only apply on stratum 2. This allows Giegerich to make a crude distinction between productive and unproductive affixation. In contrast, by providing all affixes with subcategorisation frames, and having no theory of strata, Benua's model makes no claims about morphological productivity. Benua argues this is not as a problem as productivity is nothing to do with phonology (Benua, 1997: 208).

frame is to specify whether the affix triggers an OO_1- or OO_2-correspondence relationship (Benua, 1997: 30). Because OO-correspondence relationships are linked to morphological subcategorisation frames, correspondence relationships can only apply between words that are linked by affixation, and then only to words that are separated by a single affix, e.g. *subliminal*, *subliminality* (Benua, 1997: 30, 31). Therefore, in TCT, just as under the phonological cycle, paradigmatic identity is evaluated in a 'strictly local' way. In TCT, *orìginálity* has the stress pattern it does because it preserves the pattern of its strictly local neighbour *oríginal*; it cannot copy the stress pattern of more-deeply embedded ***ó**rigin* to give ***òriginálity*** because there is no affixation process that can connect *origin* and *originality* directly (i.e. there is no suffix *-*ality*). Benua's theory also rules out identity relations between words linked by the same affix (cp. Kenstowicz, 1996: 382; Burzio, 2000, 2005a, b – see §7.3), e.g. *sanity*, *obesity*, as these are not linked by morphological derivation (Benua, 1997: 32).

So far, we have seen how TCT captures one aspect of the phonological cycle – strictly local identity relationships. Strict locality is only achievable in TCT by the stipulation that OO-correspondence relationships are linked to subcategorisation frames – it is not the null hypothesis of OO-correspondence. The second key prediction made by the phonological cycle is that morphophonological relationships will be asymmetrical: embedding words will preserve the stress of their embedded words, but not vice-versa as in ****ò**riginálity* (the stress pattern expected without preservation from *oríginal*) → ****ó**riginal* or ****ò**rigínal*. How the asymmetry of the cycle is enforced in TCT is now discussed.

7.3.1.2 'Priority of the Base'

The linear evaluation of paradigms in TCT that was outlined in §7.3.1.1 only ensures that paradigmatic identity is strictly local; it does not predict in which direction paradigmatic identity will apply. Building on work on truncation from Benua (1995), Benua (1997) argues that identity between a base and its derived word is strictly

---

However, Benua's theory aspires to model morphophonological interactions, so this excuse is not sufficient. As we will see in chapter 8, the model of Stratal OT proposed there, like Giegerich's model, makes crude distinctions about productivity which are surely relevant to a model of morphophonology.

asymmetrical, just as under the phonological cycle: the derived word may reflect the phonological identity of the base, but not vice-versa. Benua calls this generalisation 'Priority of the Base', and reinforces it in TCT through a stipulated process of 'Recursive Evaluation'.

Recursive Evaluation involves duplicating the constraint hierarchy for each separate OO-correspondence relation between a pair of strictly local words (= each 'subparadigm'), e.g. *origin* and *originality*, or *original* and *originality*. The less complex a word is, the greater priority its recursion has. This means that any candidate eliminated in the recursion belonging to a less complex word will automatically be ruled out in the recursion of a more complex word, even if it would otherwise win in that particular tableau. This process is exemplified in (6), where grey shading indicates a candidate that has been blocked from winning in later recursions by losing on an earlier recursion. In each particular recursion, the same output candidate may occur more than once; this is so that, over the entire evaluation, various possible combinations of output candidates are evaluated.

(6) Recursive Evaluation to preserve stress in Benua (1997)

NONFINAL = the final syllable is extrametrical

ALIGN-R = main stress is on the rightmost syllable of the word

OO_1-ANCHOR = edges and heads of feet correspond

ALIGN-L = every prosodic word begins with a foot

Recursion (A):[18]

Input: /*origin*/	NON FINAL	ALIGN-R	OO_1-ANCHOR	ALIGN-L	>>
a. o(rí.gin)	*!	*		*	
b. (ó.ri)gin		**			
c. (ó.ri)gin		**			
☞d. (ó.ri)gin		**			

[18] OO-correspondence is vacuously satisfied in this minimally complex word: the 'Priority of the Base' generalisation means that there is no output to correspond to, i.e. *origin* cannot look ahead to *oríginal*.

Recursion (B):

Input: /*origin-al*/ Base: *(ó.ri)gin*	NON FINAL	ALIGN-R	OO_1-ANCHOR	ALIGN-L	>>
a′. o(rí.gi)nal		**		*	
b′.(ó.ri)ginal		***!			
☞c′.o(rí.gi) nal		**	*		
☞d′.o(rí.gi)nal		**	*		

Recursion (C):

Input: /*origin-al-ity*/ Base: *o(rí.gi)nal*	NON FINAL	ALIGN-R	OO_1-ANCHOR	ALIGN-L	>>
a″.o(rì.gi)(ná.li)ty		**	*	*	
b″.(ò.ri)gi(ná.li)ty		**	*		
c″.(ò.ri)gi(ná.li)ty		**	**!		
☞d″.o(rì.gi)(ná.li)ty		**	*	*	

In (6), candidate b″**òriginálity* would win on Recursion (C) had its candidate paradigm not already been ruled out by the fatal violation of ALIGN-R by candidate b′. **óriginal* on Recursion (B).

Recursive Evaluation can clearly enforce the asymmetrical identity between words that is required by stress preservation, but it is very much a stipulation upon the null hypothesis of OO-correspondence – symmetrical identity (Bermúdez-Otero, 1999: 108, 113-121, 143-148; 2002). Benua's argument for Recursive Evaluation over a mechanism like the phonological cycle, where, in the latter, serialism comes for free, is on the grounds that Recursive Evaluation respects the parallelism of classical OT: "[t]here is no sense in which the less complex word is derived first" (Benua, 1997: 12).

In the sense that the outputs which are argued to be in an OO-correspondence relationship in TCT are available to the grammar at the same moment in time, Benua is correct in arguing Recursive Evaluation to be parallel. However, as we will now see in §7.3.1.3, the notions of 'parallelism' and 'serialism' can be interpreted in more than one way. It will be argued that Recursive Evaluation may be interpreted as a form of serialism, and that this form of serialism is also available to overtly serial models like Stratal OT. Without the advantage of parallelism, and with the stipulation of Recursive Evaluation, we will see that the case for TCT is weakened.

7.3.1.3 The notion 'serialism'

As we saw in §7.3.1.2, Recursive Evaluation means that OO-correspondence relationships are only enforced asymmetrically in TCT, in contrast to the null hypothesis of OO-correspondence. We also saw that, according to Benua, this asymmetry is achieved while still preserving the 'parallel' nature of classical OT: Recursive Evaluation means that all the words in an extended paradigm like *órigin*, *oríginal*, *orìginálity* can be processed simultaneously. The ranking of recursions as in Recursion (A) >> Recursion (B) >> Recursion (C) ensures asymmetrical identity relationships, without requiring Recursion A to take place before Recursion B, etc., in real time.

However, 'serialism' can be interpreted in more than one way. Benua's argument for TCT's 'parallelism' only holds if we assume 'serialism' to absolutely imply temporal precedence, rather than just an asymmetrical relationship. Clearly, TCT would be serial in the latter sense. If Benua's model is serial in terms of its asymmetry, if not temporal ordering, we have to consider whether the overtly serial models Benua opposes cannot, in fact, be interpreted as serial in terms of a-temporal asymmetry as well. If so, then it has to be seriously considered whether the stipulation of Recursive Evaluation upon OO-correspondence is warranted: the same type of serialism would come for free in Stratal OT.

Bermúdez-Otero (in preparation), following arguments by Marr (1982) and Arbib (1987), argues that we can differentiate between 'functional serialism' and 'algorithmic serialism':

(7) Functional and algorithmic serialism (Bermúdez-Otero, in preparation)

(a) Functional serialism: A linguistic theory is serialist at the functional level if it postulates dimensions of representation that are asymmetrically related, in the sense that representations in one dimension are a function of representations in another dimension, but not vice versa.
(b) Algorithmic serialism: [A] linguistic theory is serialist if it specifies a processing algorithm involving a sequence of temporally ordered computational operations.

So, even if, as Benua's appeal to OT's parallelism implies, TCT is not algorithmically serialist – the evaluations of all words in an extended paradigm are performed at the same time – TCT is still serial in the sense that it requires asymmetric OO-correspondence relationships: functional serialism.

The question is, therefore, whether a self-professed serial model, such as Stratal OT, can also be interpreted in terms of functional serialism. Bermúdez-Otero (in preparation) argues that it can (see also Orgun, 1996a: 17; Bermúdez-Otero, 1999: 77-79). Bermúdez-Otero shows that the phonology of morphologically complex words is determined by analogous phonological functions in Stratal OT and TCT. These are given in (8):

(8) Phonology of complex words as composite functions in TCT and Stratal OT (Bermúdez-Otero, in preparation)

(a) TCT

P = phonological function; *o* = output; *i* = input

Phonology of base output: $o_{base} = P_{IO}(i_{base})$
Phonology of derived output: $o_{derived} = P_{IO+OO}(i_{derived}, o_{base})$
therefore Phonology of derived output: $o_{derived} = P_{IO+OO}(i_{derived}, P_{IO}(i_{base}))$

(b) Stratal OT

P = phonological function; *u* = underlying representation; *s* = surface representation; *a*, *b* = phonological domains; *x*, *y* = morphemes

Underlying representation of base: $u = [_a\, x]$
Surface form of base: $s = P_a(x)$
Underlying representation of derived form: $u = [_b[_a\, x]\, y]$
therefore Surface form of derived form: $s = P_b(P_a(x), y)$

The vital point to be noted from (8) is that the phonology of the derived word is a "composite function" (Bermúdez-Otero, in preparation) in both TCT and Stratal OT: the phonological function of the derived form in TCT, $P_{IO+OO}(i_{derived}, P_{IO}(i_{base}))$, contains the phonological function of the base $P_{IO}(i_{base})$; similarly, the phonological function of the derived form in Stratal OT, $P_b(P_a(x), y)$, contains that of the base, $P_a(x)$. As argued by Bermúdez-Otero, if the phonological function for the derived and base forms can be processed simultaneously in TCT, there seems to be no reason why this cannot extend to Stratal OT: the two theories propose analogous composite functions. If the human brain can process the functions $P_a(x)$ and $P_b(P_a(x), y)$ simultaneously, as is argued by TCT, then there seems to be no reason why it cannot also compute the two functions of Stratal OT's composite function, $P_{IO}(i_{base})$ and $P_{IO+OO}(i_{derived}, P_{IO}(i_{base}))$, in parallel.

In sum, there is nothing to suggest that TCT enjoys any advantage in terms of parallelism over an overtly serial model like Stratal OT. This is a very important result: one of Benua's main arguments for TCT, and hence the reason for the stipulation of Recursive Evaluation upon OO-correspondence, is that TCT can maintain OT's strong claim to parallelism. If TCT is, in fact, serial, then Recursive Evaluation is an unnecessary stipulation: serialism comes for free in overtly serial theories. For example, as we see in §7.4 next, Stratal OT enforces identity effects through Input-Output Correspondence constraints. The identity enforced by Input-Output Correspondence constraints is inherently serial (Kiparsky, 1998a: 8), cp. OO-correspondence constraints, as the OT computation proceeds only ever from input to output, not vice-versa (see figures 1 and 2, above).

TCT's serialism is a damning blow for the theory, given that serialism comes for free in Stratal OT. However, the argument against TCT and for Stratal OT is not this simple: TCT's stipulation of serialism may be justifiable if Stratal OT is equally or more stipulative in other respects. Stratal OT is discussed further in §7.4, where it is shown that the balance of arguments is in favour of Stratal OT.

7.4 Paradigmatic opacity, route 2: Stratal OT

As outlined in §1.4.1, Stratal OT is a hybrid model of OT that broadly combines the insights of Lexical Phonology and Morphology with OT's theory of parallel constraint evaluation. In contrast to the classic monostratal theory of OT, and in common with theories of Lexical Phonology, Stratal OT consists of not a single phonology, but several serially ordered cophonologies (figure 3):

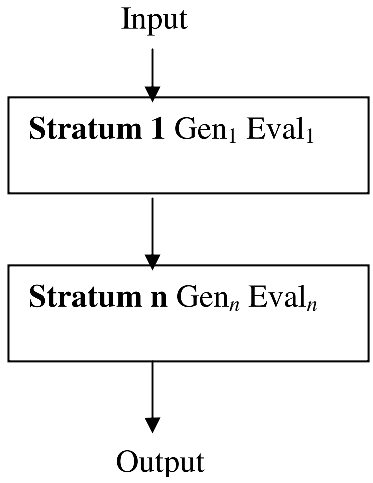


Figure 3: Stratal OT (Kager, 1999: 382)

Each phonological level corresponds to a morphological category, e.g. Stem, Word, but, like classical OT, the phonological computation within each individual cophonology takes the form of output-orientated constraint evaluation.

In contrast to TCT, Stratal OT permits faithfulness between Input and Output ('IO-correspondence'), but not between outputs. IO-correspondence is inherently serial: OT derivations proceed only ever from input to output, not vice-versa. OO-correspondence should not be required in Stratal OT: paradigmatic opacity is achieved through the serial ordering of cophonologies which are linked by IO-correspondence, and so another means of enforcing paradigmatic opacity, OO-correspondence, should be superfluous.

In the models of Stratal OT proposed by both Bermúdez-Otero (forthcoming) and Kiparsky (1998a, 2000, 2003a, b, forthcoming), there are three serially ordered cophonologies, corresponding to the morphosyntactic categories Stem, Word and Phrase respectively. Opacity is achieved by the output of an earlier constraint hierarchy (e.g. the stem-level cophonology) being only partially overwritten in an evaluation by a later constraint hierarchy (e.g. the word- or phrase-level cophonology); IO-correspondence can prevent the outputs of previous levels being completely overwritten.

Serialism in Stratal OT is intrinsic as a result of the intrinsic morphosyntactic constituency of words and phrases (Kiparsky, 2000, 2003a; Bermúdez-Otero, in preparation). Words and phrases consist of morphosyntactic categories which are like layers of an onion: Root is the innermost, then Stem, then Word, then Phrase. Because cophonologies correspond to this 'onion-like' morphosyntactic structure, they must necessarily apply one after the other: the phonology which corresponds to an outer morphosyntactic domain must apply after the phonology corresponding to an inner domain.[19]

An analysis of English strong stress preservation can exemplify how paradigmatic opacity is achieved in Stratal OT. Under strong preservation, the main stress of an embedded word is preserved as a main stress in the embedding word: e.g. ***ób**vious* → ***ób**viousness*. In the word *óbviousness*, primary stress application has misapplied: English primary stress can apply within a three-syllable window at the right edge of the word (in nouns), but primary stress is on the preantepenult in *óbviousness*. As shown in (9), strong preservation is dealt with in Benua's (1997) model of TCT by ranking the OO-correspondence constraint associated with class 2 affixes like *-ness* – OO_2-ANCHOR – above the constraints that ensures primary stress is within the three syllable window at the right edge of a word, ALIGN-R (ALIGN ((Hd)PrWd, R, PrWd, R)):

(9) Strong preservation in TCT (Benua, 1997: 176)[20]

NONFIN = final syllable is extrametrical

OO_2-ANCHOR = stress of derived word corresponds with that of base

ALIGN-R = main stress is on the rightmost syllable of the word

[19] This serialism therefore predicts that phrase-level properties cannot be triggered by morphological structure. Advocates of OO-correspondence have attempted to argue that there are phrase-level processes which require morphological conditioning, but there are counter-arguments to this in Stratal OT, e.g. Bermúdez-Otero & McMahon's (2006: §3.6) analysis of the Withgott effect.

[20] Benua (1997) shows the second syllable of *óbvious* as not belonging to a foot. Benua does not deal with the ranking of PARSE-SYLL ('all syllables must belong to a foot'), which, here, could force the second syllable of *óbvious* to belong to the preceding foot. As its stands, both the parsing shown by Benua, and the alternative with the second syllable footed, satisfy the formulation of FTBIN used by Benua – 'feet are binary on a moraic or syllabic analysis'.

Recursion (A)

Input: /*obvious*/	NON FINAL	OO_2-ANCHOR	ALIGN-R	>>
a. ob(ví.ous)	*!		*	
b. (ób)vi.ous			**	
☞c. (ób)vi.ous			**	

Recursion (B)

Input: /*obvious-ness*/ Base: *(ób.vi)ous*	NON FINAL	OO_2-ANCHOR	ALIGN-R	>>
a. ob(ví.ous)ness			**	
b. ob(ví.ous)ness		*!	**	
☞c. (ób.)vi.ous.ness			***	

In Stratal OT, the asymmetric relationship enforced by Recursive Evaluation in Benua (1997) is instead ensured by passes through the stem- and word-level cophonologies, in that order. IO-correspondence preserves the output of a previous level, resulting in the misapplication of primary stress in *óbviousness*:

(10) Strong preservation in Stratal OT[21]

IDENT-IO(STRESS) = stress of output corresponds with that of input

Stem level			
Input: /*obvious*/	NON FINAL	ALIGN-R	IDENT-IO(Stress)
a. ob(ví.ous)	*!	*	
☞b. (ób)vi.ous		**	

⇓

[21] IDENT-IO(Stress) is an approximate metrical IO-correspondence constraint which suffices for this analysis. More precise IO-correspondence constraints are shown to be necessary for the analysis of weak stress preservation in chapters 8 and 9.

Word level			
Input from Stem level: *ób.vi.ous-ness*	NON FINAL	IDENT-IO(Stress)	ALIGN-R
a. ob(ví.ous)ness		*!	**
☞b. (ób)vi.ous.ness			***

The winning candidate on the Stem level – b. *óbvious* – is the input to the Word level, where it is suffixed with the word-level suffix -*ness*. Faithfulness to stress specified on the word-level input *óbvious+ness* is ranked higher than ALIGN-R, and so the misapplying stress pattern *óbviousness* is the output of the Word level. Strong preservation will not occur on the Stem level. On the Stem level, stress shifting occurs, so that primary stress is always close to the right edge of the word – weak preservation. This is ensured by the different ranking of ALIGN-R and IDENT-IO(Stress) on the Stem level as compared to the Word level.

By consisting of several serially ordered cophonologies, Stratal OT is at odds with one of the assumed defining characteristics of OT: strong parallelism. However, the incompatibility of serialism and OT may be a perceived rather than actual truth. Cohn & McCarthy (1994: f.n. 2), Clements (2000) and Booij (1996) all make arguments to the effect that it is the theory of parallel, surface-orientated constraint interaction that is OT's valuable contribution to phonological theory; the strong emphasis upon parallel evaluation of all candidates is altogether more dispensable. There is no doubt that any version of serialism proposed in OT will be more constrained than that of rule-based theories. The nature of parallel constraint interaction in OT absolutely places some limit upon opacity by excluding the possibility of stipulative rule orderings that can generate many abstract intermediate representations (Booij, 1996; Bermúdez-Otero, 1999: 93-4; Kiparsky, 2000: 352).[22] However, OT does not predict the nature of the morphophonology interface or the greater organisation of the grammar in general (Bermúdez-Otero, in preparation): parallel constraint interaction is merely a theory of phonological computation. The interaction of OT phonology with morphology may therefore indeed be serial.

[22] Although precisely the sort of stipulative opacity achievable through rule ordering is seen in OT in Sympathy Theory: the sympathy candidate is an ad hoc representation, just like the intermediate representations that come with rule-ordering (Bermúdez-Otero, 1999: 149; Kiparsky, 2003a: 263). As shown in §7.4.2, Stratal OT has no need for Sympathy Theory.

A model of OT which incorporates the phonological cycle is not only entirely feasible – it is potentially desirable. Critics of OT have cited OT's vague approach to morphology as one of the theory's weaknesses (as noted by McMahon (2000: 44-7) and Hurrell (2001: 34)), a similar problem occurred in the early days of Lexical Phonology and Morphology, e.g. Halle & Mohanan (1985)); this is an issue Stratal OT takes to task. The hybrid model advocated by Stratal OT potentially offers the best of both worlds. Lexical Phonology and OT complement one another by filling in the gaps in the other theory (Kiparsky, 2003a: 257, 2007a: 10). On the one hand, OT's theory of constraint interaction places limits upon unprincipled, abstract intermediate representations, very desirable where stress is concerned (§7.1): in Stratal OT, the only intermediate representations permitted are those predicted by morphosyntactic structure. On the other hand, Lexical Phonology's theory of strata complements OT by providing it with a model of morphophonology interaction, and hence a mechanism of phonological opacity.

Stratal OT still has a competitor when it comes to modelling paradigmatic opacity in OT: if Stratal OT is not damned by its serialism, then neither is Benua's model of OO-correspondence, TCT, which was shown to be serial in §7.3.1.3. It will already have been observed that there are some fundamental differences between Stratal OT and TCT: a plurality of OO-correspondence constraints (TCT) versus just IO-correspondence (Stratal OT); a single fixed constraint ranking (TCT) versus constraint re-ranking (Stratal OT). In the rest of this section, Benua's (1997) arguments against Stratal OT are considered, and shown to be either unjustified, or met with an equal issue in TCT.

7.4.1 OO-correspondence versus cophonologies

In monostratal OT, just a single constraint ranking is proposed for a language. This is because paradigmatic opacity effects can be handled by introducing a new set of constraints: OO-correspondence constraints. As we have seen, this is not the case in Stratal OT: Stratal OT argues for only IO-correspondence constraints, and therefore paradigmatic identity must be allowed for by multiple different constraint rankings or 'cophonologies'.

In (10), we saw that it was necessary to rerank faithfulness and markedness constraints between the stem- and word-level cophonologies in order to account for the misapplication of primary stress seen with word-level suffixations (strong preservation), but not with stem-level suffixations (only weak preservation occurs on the Stem level). Benua argues that this type of re-ranking is a weakness in Stratal OT: it predicts too much typological freedom within a single language, as the predictions made by the cophonologies of any language could diverge quite considerably and improbably (Benua, 1997: 90-95).[23] In reality, this wide divergence is not observed. Stratal OT has, of course, attempted to answer this problem: e.g. Bermúdez-Otero (1999: 98-104, 108, 153-4; 2003) argues that the degree of divergence between cophonologies is limited both by language acquisition and historical factors. Nevertheless, both language acquisition and history are external constraints upon Stratal OT, and so Benua seems reasonable in arguing that TCT is intrinsically more falsifiable than Stratal OT in this respect: in TCT, the possibility of reranking between strata is obviated because the theory is monostratal.

However, it appears that the typological freedom permitted by Stratal OT is, in fact, necessary. In TCT, only different rankings of faithfulness constraints are required to account for class-1 versus class-2 behaviour: the ranking of markedness constraints is consistent throughout the single phonology. Benua observes that to accommodate this insight into Stratal OT would require a costly stipulation: namely, that only faithfulness and not markedness constraints can rerank between strata. Fortunately, Bermúdez-Otero (1999) shows Benua's argument to be incorrect, giving his own examples of synchronic markedness reversal (e.g. West Germanic Gemination; 1999: 186) which require markedness-constraint reranking.

TCT also suffers its own problems in terms of restrictiveness. As well as stipulating serialism (§7.3.1.3), TCT introduces a whole new constraint type into CON – OO-correspondence. Benua (1997) defends OO-correspondence from the perspective of constraint grounding, arguing that OO-correspondence explicitly recognises the requirement for paradigmatic identity between morphologically related words. It is understandable why Benua might make this argument for OO-correspondence in terms of grounding: as noted by Bermúdez-Otero (1999: 152), OT

[23] Benua exemplifies this with a discussion of "'backwards application'" of nasalisation in fictitious languages.

constraints are always ideally grounded in some way, and OO-correspondence explicitly recognises the functional role phonological similarity plays in phonological recognition. But there is also the flip side of Benua's argument to be considered. Bermúdez-Otero (1999: 107) notes that, while IO-correspondence is required anyway in any theory of OT, OO-correspondence is introduced especially to enforce paradigmatic identity, and, as evident from the discussion so far in this chapter, OO-correspondence has the potential to be far more powerful than actually seems to be required, being able to enforce faithfulness between **any** outputs. OO-correspondence models precisely and only the type of paradigmatic identity that comes for free under IO-correspondence, so that to introduce OO-correspondence too is formally uneconomical.

TCT also contains no intrinsic limits on how OO-correspondence constraints must rank. We have seen that Benua proposes two families of OO-correspondence constraints that predict the two degrees of opacity observed in English: OO_1, and OO_2. In this way, Benua can capture the two distinct sets of effects that follow from class-1 and class-2 suffixation in English: notably, the greater faithfulness of words affixed with class-2 suffixes than class-1 suffixes. However, given the potential freedom of faithfulness available under a theory of OO-correspondence, that just these two meta-rankings occur is odd. In TCT, OO_1 and OO_2 are proposed on phonological grounds alone – it is pure stipulation that ensures we do not also have OO_3 and OO_4 meta-hierarchies, and so on. No non-phonological factors converge with the phonological observation that precisely two meta-hierarchies are required.

Benua's need to stipulate two phonological classes is symptomatic of a more general problem observed in non-stratal models. Although work like Zamma (2005) rightly points out that there is further sub-variation in morphophonological behaviour between the two major types of affixation in English, there is a consensus, even between those who reject a stratal approach (e.g. Burzio, 1999, 2000, 2005b), that there are just two overarching types of behaviour. This does not fall out of TCT, where as many phonological classes may be proposed as types of phonological behaviour observed. In TCT, only external factors, e.g. learnability, may constrain the number of phonological patterns observed in a language (Benua, 1997: 229).

In contrast, in Stratal OT, the clustering of morphophonological behaviour into just two over-arching types converges with morphological predictions, and is therefore highly restrictive. The number of different constraint rankings permitted is constrained by intrinsic morphological and prosodic constituency (Kiparsky, 2000: 351). Words and phrases intrinsically consist of grammatical categories like stems, words and phrases. These categories correspond to phonological domains, and there can be no phonological domain which does not correspond to a morphosyntactic category. In sum, a stratal model "allows the morphology to tell the learner what phonological behaviour to expect" (Kiparsky, 2000: 362).[24] With three morphosyntactically motivated levels in Stratal OT – Stem, Word and Phrase – precisely two degrees of opacity are predicted: exactly what is attested (Kiparsky, 2003a: 263). Hurrell (2001) makes a similar case for a stratal model:

> [W]ithout a more clearly defined approach to morphology, there will be no apparent limit on the number of different sets of rankings the theory [i.e. OT] can admit – a problem analogous to the multiplying-levels problem in 1980's Lexical Phonology (Hurrell, 2001: 35).

TCT can be rightly said to suffer in this respect.

So far, then, TCT's arguments against Stratal OT do not stand: the unconstrained constraint reranking permitted by Stratal OT is, in fact, necessary, whereas TCT's introduction of OO-correspondence constraints is not. Stratal OT can also offer principled limits upon the depth of phonological opacity in terms of its theory of strata, whereas TCT relies upon some external constraint. These advantages of Stratal OT over TCT have to be added to our observation from §7.3.1.3 that serialism comes for free in Stratal OT, but must be stipulated in TCT. In short, the case so far is very much in favour of Stratal OT. Now, in §7.4.1.1, a particular situation is considered in which TCT makes the wrong predictions with respect to English stress preservation.

---

[24] Kiparsky (2000: 362) proposes that 'Stem' and 'Word' are part of UG, in the same way that 'noun' and 'verb' are. Bermúdez-Otero (in preparation) speculates that Stratal OT's morphophonological levels are not an innate part of UG, but may be formulated in language acquisition (although Bermúdez-Otero still proposes that the three morphophonological levels are characteristic of all grammars).

7.4.1.1 Exceptional and preserving stress

As pointed out by Kiparsky (2007a), it has been noted in the literature on cyclicity that the evidence for the cyclic application of phonological processes, e.g. stress preservation, occurs in the phonological contexts where the relevant feature is lexically contrastive, i.e. lexical exceptions. Following Bermúdez-Otero (2007b, in preparation), this will be referred to as 'Chung's Generalisation' – Chung (1983: 63) observes the connection between cyclic and lexically contrastive properties (see also Kiparsky, 2007a: 20). Chung's generalisation is defined by Bermúdez-Otero (2007b) as in (11):

(11) Chung's Generalisation (Bermúdez-Otero, 2007b: 13)

If a stem-level phonological generalization displays cyclic misapplication, then it also has lexical exceptions.

Chung's Generalisation has been implicit in Pater's (1995, 2000) work on preservation in OT. Pater (1995, 2000) shows that what we regard as cyclic pre-tonic vowel-quality preservation – in his argument, stress preservation – occurs in the same phonological contexts as exceptionally full vowels occur in words which are not candidates for preservation, e.g. *cond*[ɛ́]*nse* → *cònd*[ɛ]*nsátion* and *ìnc*[æ]*ntátion* (← **incánt(ate)*), cp. *cómp*[ə]*nsàte* → *còmp*[ə]*nsátion*. Very interestingly, Pater (1995: 22, f.n. 11) makes the same observation for the left-edge stress-preservation context: preservation is evident in the same phonological context in *oríginal* → *orìginálity* as exceptional stress is in *Epàminóndas* (cp. *àbracadábra*). (N.B. We did explore alternative explanations for the secondary stress of *Epàminóndas* in §5.2.2, but no totally satisfactory solution was found, indicating that some form of lexical marking is required.) In the case of both pre-tonic vowel-quality preservation and left-edge stress preservation, there is evidence of preservation – traditionally a cyclic process – in the precise phonological contexts where the phonological feature in question is lexically contrastive.

Kiparsky (2007a) notes that Chung's Generalisation falls naturally out of Stratal OT. First, consider the exceptional stress of *Epàminóndas*. To ensure that *Epàminóndas* has stress on its second syllable, exceptional stress present on this syllable in the input must be shielded by a faithfulness constraint that outranks any conflicting markedness constraints: FAITHFULNESS>>MARKEDNESS preserves underlying lexical contrasts in OT (Kager, 1999: 4-5). Stratal OT permits only a single type of faithfulness constraint – IO-correspondence. Both exceptional and preserving stress must therefore be ensured by the effects of IO-correspondence. This consequently predicts that preserving stress occurs in the phonological context where stress is lexically contrastive (Kiparsky, 2007a: 27), as shown in (12):

(12) Exceptional and preserving stress in Stratal OT

Input: /*phe(nó.me)non-ology*/	IDENT-IO (stress)	ALIGN-L
☞a. phe(nò.me)(nó.lo)gy		*
b. (phè.no)me(nó.lo)gy	*!	
Input: /*E(pá.mi)nondas*/		
☞ a. E(pà.mi)(non)das		*
b. (È.pa)mi(nón)das	*!	

In contrast, Kiparsky (2007a) points out that any theory which uses OO-correspondence constraints in addition to IO-correspondence makes the wrong predictions. In such a theory, stress preservation will be assigned to OO-correspondence, and exceptional stress to IO-correspondence. Consequently, the link between lexically contrastive and preserved stress is lost: OO-correspondence and IO-correspondence constraints can be ranked differently, so that preservation can occur without exceptional stress occurring in the same context, and vice-versa, as shown in (13):[25]

[25] This characteristic of OO-correspondence is actually argued to be one of the theory's advantages by Alber (1998: 123). Alber is clearly unaware of the present argument from English.

(13) Exceptional and preserving stress in OO-correspondence

(a) Exceptional stress but no preservation

Input: /*imagine-ation*/ Base: *i(má.gi)ne*	IDENT-IO(stress)	ALIGN-L	OO-FAITH (Stress)
☞a. (ì.ma)gi(ná)<tion>			*
b. i(mà.gi)(ná)<tion>		*!	
Input: /E(pá.mi)nondas/			
a. (È.pa)mi(nón)<das>	*!		
☞ b. E(pà.mi)(nón)<das>		*	

(b) Preservation but no exceptional stress

Input: /*imagine-ation*/ Base: *i(má.gi)ne*	OO-FAITH(Stress)	ALIGN-L	IDENT-IO (stress)
a. (ì.ma)gi(ná)<tion>	*!		
☞ b. i(mà.gi)(ná)<tion>		*	
Input: /*E(pá.mi)nondas*/			
☞ a. (È.pa)mi(nón)<das>			*
b. E(pà.mi)(nón)<das>		*!	

In conclusion, because it uses both OO-correspondence and IO-correspondence constraints, TCT misses a vital relationship between exceptional, contrastive stress and systematically preserved stress; Stratal OT suffers no such problem. In chapter 8, some further implications of Chung's Generalisation for the handling of weak stress preservation in Stratal OT are considered.

7.4.2 Non-paradigmatic opacity

So far in this chapter, only paradigmatic opacity has been addressed. Instances of non-paradigmatic opacity – opaque phonology not attributable to another output – also occur. Non-paradigmatic opacity cannot be handled by OO-correspondence, and so Sympathy Theory (McCarthy, 1998) has been proposed to handle non-paradigmatic opacity in monostratal OT. In contrast, in Stratal OT, both paradigmatic and non-paradigmatic opacity can be handled by the same mechanisms (Bermúdez-Otero, in preparation).[26]

The following example of Canadian Diphthong Raising (Bermúdez-Otero, 2003, in preparation) shows how paradigmatic and non-paradigmatic opacity are handled alike in Stratal OT. In Canadian English, the diphthongs /aɪ/ and /aʊ/ raise to give [əi] and [ʌʊ] respectively when followed by a voiceless obstruent in the same prosodic word which does not belong to a metrically stronger syllable, as shown in (14):

(14) Canadian Diphthong Raising (Bermúdez-Otero, 2003)

[rəit]	*write*	cf.	[raɪd]	*ride*
[nə́itrèɪt]	*nitrate*	cf.	[sàɪtéɪʃn]	*citation*

In Canadian English, there is also phrase-level flapping of /t/ and /d/ which occurs when either is lax (not foot-initial), preceded by a vowel or /r/, and followed by a vowel (Bermúdez-Otero, 2003: 8). This results in phrase-level alternations, as shown in (15):

(15) Phrase-level flapping in Canadian English (Bermúdez-Otero, 2003)

a. [fæɾər]	*fatter*	cf. [fæt]	*fat*
b. [mæɾər]	*madder*	cf. [mæd]	*mad*
c. [hi hɪɾ æn]	*he hit Ann*	cf. [hɪt]	*hit*
d. [hi hɪɾ æn]	*he hid Ann*	cf. [hɪd]	*hid*

(15c, d) show flapping is Phrase level as it occurs across a word boundary. The final /t, d/ of *hit* (15c) and *hid* (15d) are lax, as they are not foot-initial at the Word level.

[26] It is likely that descriptive economy is not the only advantage Stratal OT gains by not requiring Sympathy Theory. Sympathy Theory is now infamous for its various problems, in particular, the great increase in the complexity of the grammar by extending phonological faithfulness to candidate-to-candidate relationships, and learnability (Kager, 1999: 392; Kiparsky, 2000, 2003a: 263; Bermúdez-Otero, 2003).

Famously, Canadian Diphthong Raising overapplies before flapped /t/ (i.e. Diphthong Raising applies in contexts where it is not followed by a voiceless obstruent) (Bermúdez-Otero, in preparation). Diphthong Raising is opaque in both derived **and** non-derived words, as shown in (16):

(16) Overapplication of Canadian Diphthong Raising

a.	/raɪt/	[rəiɾər] *writer*	cf. [rəit] *write*
b.	/maɪtər/	[məiɾər] *mitre*	

In both *writer* and *mitre*, the conditioning environment for Diphthong Raising – the following /t/ present in the underlying representation – is overwritten by phrase-level flapping. Crucially, the flapping environment is either present in the underlying representation, as in *mitre*, or is the result of a morphological process, as in *writer* (*write* + *-er*).

Writer can be viewed as an instance of paradigmatic opacity – *writer* has the diphthong [əi] opaque-ly because this diphthong occurs transparently in morphologically related *write*. *Mitre* has to be an instance of non-paradigmatic opacity: there is no related word which can account for overapplication of Diphthong Raising in *mitre*. In Stratal OT, the paradigmatic opacity of Diphthong Raising in *writer* and the non-paradigmatic opacity of this process in *mitre* both follow from the model's stratal structure, as shown in (17) (from Bermúdez-Otero, 2003):

(17) Diphthong Raising overapplies as a consequence of level ordering

Domain structure $[_{PL}[_{WL}[_{SL}$*write*$]$*er*$]]$ $[_{PL}[_{WL}[_{SL}$*mitre*$]]]$

(SL = Stem Level

WL = Word Level

PL = Phrase Level)

Underlying Representation	/raɪt/	/maɪtər/
Stem level – Raising	[rəit]	[məitər]
Word level	[rəitər]	[məitər]
Phrase level – Flapping	[rəirər]	[məirər]

Whether or not a morphological item undergoes any "overt" morphological process on a particular level – e.g. affixation of *write+er* on the Word level – or not, as in *mitre*, it must still pass through the level and therefore also its phonology (Bermúdez-Otero, 1999: 140). It is through the "covert" morphological operation of simply passing through a particular level that non-paradigmatic opacity effects can occur in Stratal OT (Bermúdez-Otero, 1999: 140): *mitre* undergoes no overt morphology, but must still pass through all levels.

In monostratal OT, both *writer* and *mitre* cannot be handled the same way. OO-correspondence can handle overapplication of Diphthong Raising in morphologically complex *writer*, but Sympathy Theory is required to account for the overapplication of Diphthong Raising in simplex *mitre*.[27] A suggestion of how this could be done is given in (18):[28]

(18) Sympathy and OO-correspondence (based on Bermúdez-Otero, in preparation)

CLEARDIPH = distance between two elements of a diphthong is maximal

IO-IDENT(seg) = output segments must be the same as that of the input segments

OO-IDENT(Diph) = derived output diphthong must be same as that of base output

IDENT-❀O(Diph) = derived output diphthong must be same as that of sympathy candidate

[27] It does not appear to be the case that Sympathy Theory can handle both types of phonological opacity, both paradigmatic and non-paradigmatic: in accounting for the feeding Duke-Of-York derivation seen in the interaction of schwa epenthesis, schwa syncope, and spirantisation in Tiberian Hebrew, McCarthy (2003: §3.2) employs OO-correspondence (see also Bermúdez-Otero, in preparation).

[28] Bermúdez-Otero (in preparation) suggests the approximate rankings and the markedness constraint given in (18); I am responsible for the tableaux and precise faithfulness constraints used.

(a) Paradigmatic opacity: OO-correspondence and *writer*

Input: /raɪɾər/ *Base:* [rəit]	OO-IDENT(DIPH)	CLEARDIPH
a. raɪɾər	*!	
☞ b. rəiɾər		*

(b) Non-paradigmatic opacity: Sympathy Theory and *mitre*

Input: /maɪɾər/	IDENT-❀O(Diph)	CLEARDIPH	IO-IDENT(seg)
a. maɪɾər	*!		
☞ b. məiɾər		*	*
❀ c. məitər		*	**!

In (18), two totally disparate mechanisms are required to account for overapplication of Diphthong Raising in *writer* and *mitre*. Non-paradigmatic opacity in *mitre* can only be accounted for by Sympathy Theory (18b), where a sympathy candidate is chosen in which Diphthong Raising applies transparently.

Stratal OT's handling of paradigmatic and non-paradigmatic opacity is promising with respect to language acquisition, as shown by Bermúdez-Otero (2003). By handling both paradigmatic and non-paradigmatic opacity with the same mechanism, children can use the phonological alternations seen in paradigmatic opacity to discern the underlying representations of non-alternating forms: e.g. the alternation *write* [rəit] ~ *write odes* [rəiɾ oʊdz] allows the speaker to work out that the underlying representation of *mitre* [məiɾər] contains a /t/, not */ɾ/, even though *mitre* does not alternate. This is not possible in the OO-correspondence-Sympathy handling of phonological opacity: no link is made between paradigmatic and non-paradigmatic opacity.

Ironically, given monostratal OT's lack of economy when handling phonological opacity, one of the criticisms levelled by Benua (1997) against stratal models is that these models are not economical in terms of the mechanisms by which they enforce paradigmatic opacity in particular. Specifically, Benua (1997) criticises stratal models for having two mechanisms of enforcing paradigmatic opacity:

(i) strata

(ii) stratum-internal cyclicity (stratum one of most models of Lexical Phonology – see §1.2.1.2)

In contrast, in TCT, just a single mechanism brings about paradigmatic opacity: OO-correspondence. Benua's model therefore has a clear advantage over any model that employs both strata and stratum-internal cyclicity as far as paradigmatic opacity is concerned (Benua, 1997: 169). However, as I have just shown, this flaw in stratal models may be tempered by the fact that stratal models can also deal with non-paradigmatic opacity, unlike OO-correspondence. Indeed, the model of Stratal OT advocated in this thesis can go one better. As will be shown in chapter 8, the model of Stratal OT advocated here (Bermúdez-Otero & McMahon, 2006; Bermúdez-Otero, in preparation) relies upon precisely just one mechanism of enforcing paradigmatic, and, indeed, all, phonological opacity: strata. Stratum-internal cyclicity has no place in the model ('fake cyclicity' – §1.4.3). This issue is discussed at length in chapter 8; for now, it suffices to say that this rejection of stratum-internal cyclicity answers Benua's criticism.

7.4.3 Other issues: underlying representations and bound roots

In §7.4.1 and §7.4.2, two of the most significant arguments made by Benua against Stratal OT were answered: the permissiveness of constraint reranking, and the lack of economy in Stratal OT's opacity mechanisms. Here, Benua's other two arguments against stratal models are presented, and shown to be unfounded.

7.4.3.1 Access to underlying representations: [mn]-simplification

English words exhibit simplification of tautosyllabic [mn]-clusters to [m] in word-final position: *cond**emn*** [-ɛm] versus *cond**emna**tion* [-ɛmneɪ-].

Cluster-simplification overapplies in instances of class 2 suffixation, e.g. *cond**emni**ng* [-ɛmɪ-], *[-ɛmnɪ-]: cluster simplification applies even though the [mn]-cluster is not word-final. The standard argument (e.g. Kaisse & Shaw, 1985;

Mohanan, 1986), and that assumed by Benua (1997), is that [mn] must be underlying in *condemn* in order for [n] to surface in *condemnation*, even though the [n] is not apparent in the surface realisation of *condemn*.

Benua (1997: 219) argues that cluster simplification "turns up a significant problem with the serial OT grammar". Benua assumes that cluster simplification applies on stratum one in Stratal OT, as it overapplies before word-level suffixes.[29] Benua's argument is that if stratum one words are derived strictly in series, as in the Lexical Phonology's internally cyclic formulation of stratum one, then IO-correspondence would predict that we get *conde**mn*** [-m-] → **conde**mn**ation* [-m-], not *conde**mn**ation* [-mn-]. This is because *condemnation* has no means of accessing the underlying representation of *condemn* which contains the necessary [mn]-cluster.

The problem of the overapplication of [mn]-simplification in **condemnation* [-m-] does not occur in Benua's theory TCT. With the use of both IO-correspondence constraints and OO-correspondence, it is possible for *condemnation* to be simultaneously faithful to an underlying representation which contains the necessary [mn]-cluster – /kɒndɛmn-eɪʃən/ - as well as to the surface form of the word it embeds, *condemn*. This is shown in figure 4:

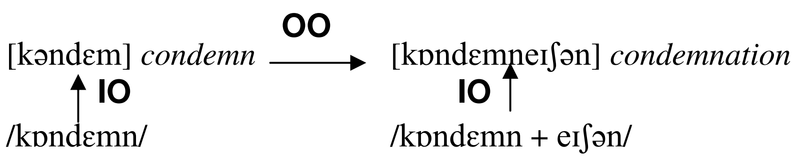


Figure 4: prevention of [mn]-simplification in TCT

As shown by Benua (1997: 199), TCT can predict the correct result *condemnation* [-mn-] under Recursive Evaluation.

Benua's criticism is fair, but means of avoiding the problem she cites in Stratal OT have been proposed by Bermúdez-Otero & McMahon (2006: §3.4). Bermúdez-Otero & McMahon, like Benua, point out that the serial derivation of

[29] In some models of LPM, it was argued that nasal-cluster simplification was a stratum two process, either applying morpheme- (not word-) finally, e.g. [[*damn*]*ing*] → [[*dam*]*ing*] (e.g. Mohanan, 1986), or by ordering all stratum two phonology before rather than after the stratum two morphology (Borowsky, 1993). Both types of approach have received criticism: see Giegerich (1999: 129-30) and Bermúdez-Otero & McMahon (2006: 395, 398). Giegerich (1999) takes an altogether different approach, proposing a stratum one process of [n]-insertion that is guided by orthography.

words like *damnation* on the highest stratum predicts the wrong result: **damnation* with just [m], not [mn]. Nevertheless, like Benua, Bermúdez-Otero & McMahon argue that [mn]-simplification must be a stem-level phonological process in a stratal model, in order to account for its overapplication in word-level forms like *damning*. Bermúdez-Otero & McMahon's way around the *condemn* [-m-] → **condemnation* [-m-] problem comes from the rejection of the stratum-internal cycle ('fake cyclicity').

As will be discussed in detail in chapter 8 (see also §1.4.3), Bermúdez-Otero & McMahon (2006) propose that stratum one is internally noncyclic, all outputs of the stratum being derived in a single pass through the stratum's phonology. All outputs of the Stem level are listed in the lexicon, and the phonological relationships between stem-level forms, traditionally captured by the phonological cycle, take the form of lexical redundancy rules. Roots are also recognised as morphological constituents, but they are argued not to constitute phonological domains.

Under Bermúdez-Otero & McMahon's formulation of stratum one, derived forms like *damnation* can have continued access to the assumed underlying representation, /dæmn/, just as in Benua's model. This can happen because there is no stratum-internal cycle. There is assumed to be a root, /dæmn/. It is well known that the suffix -*ation* can attach directly to roots, as shown by attachment to a bound root in *ovation*. Because the cyclic derivation /dæmn/ → /dæm/ → *[dæmeɪʃən] need not be adhered to – the stratum-internal cycle has been rejected – the correct form of *damnation* can be achieved by directly suffixing the root: /dæmn/ → /dæmneɪʃən/. This is diagrammatised in figure 5:

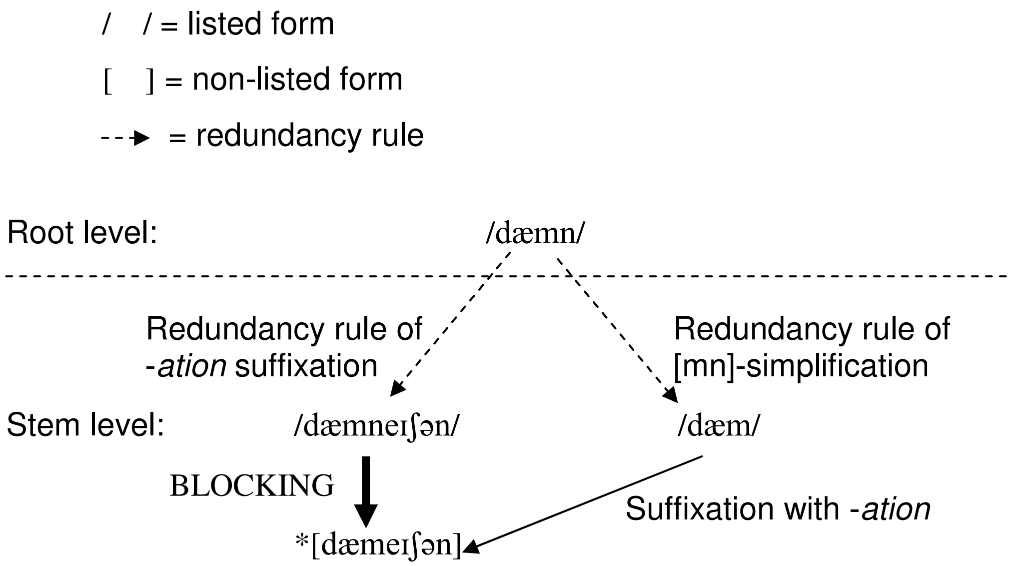


Figure 5: [mn]-simplification in Stratal OT

Also shown in figure 5 is the means by which the absence of *[dæmeɪʃən] is accounted for. Because all of the outputs of the Stem level are listed as lexical entries, blocking (Aronoff, 1976) can occur: a listed form can block another form which would perform the same function. (Blocking is discussed at length in §8.2.) In this instance, the listed form /dæmneɪʃən/ blocks the possible form *[dæmeɪʃən]: the non-occurrence of *[dæmeɪʃən] is simply an accidental gap. Importantly, the non-existence of *[dæmeɪʃən] cannot be attributed to the lack of a word-formation process that could create it: -*ation* can attach to free forms like *damn* /dæm/ as well as (bound) roots, e.g. *sum* ~ *summation*.

In conclusion, [mn]-simplification is not an argument for Benua's theory of TCT over Stratal OT: Stratal OT has proposed an alternative solution to this problem. In chapter 8, the mechanisms which are crucial to the Stratal OT analysis of [mn]-simplification – rejection of the stratum-internal cycle, listing of stem-level outputs as lexical entries, and blocking – are presented in depth. There it will be seen that these are not merely ad hoc stipulations necessary to account for [mn]-simplification, but refinements to the stratal model which offer an explanation for the probabilistic behaviour of weak stress preservation.

7.4.3.2 Bound roots

Benua argues that TCT is preferable to cyclic theories by virtue of its explicit acknowledgement that 'cyclic' effects are limited to free words:

> [I]t has been noticed since the introduction of the cycle, that cycles of rules apply only to full words, and not to smaller morphological constituents [...] Serial theories require some extra stipulation to prevent rules from cycling on bound roots [...] In TCT, the fact that bound roots are not cyclic domains follows from the basic premises of the theory (Benua, 1997: 5).

Benua is indeed correct: for example, in Giegerich's (1999) model of Lexical Phonology, it must be stipulated that bound roots must undergo some sort of overt morphological process, e.g. affixation, before being an input to the first phonological cycle (Giegerich, 1999: 111). Similarly, my understanding of Bermúdez-Otero's (in preparation) version of Stratal OT is that it is stipulated that roots coming onto the Stem level must undergo Root-to-Stem derivation before being phonological domains ("Root-to-stem derivation is stem level" – Bermúdez-Otero (in preparation)). Hence, all bound roots are stipulated to undergo suffixation before they can be phonological domains.

However, although the ethos of OO-correspondence does indeed require that paradigmatic identity only occurs between free words, the means by which this is achieved in TCT is entirely stipulative and in no way superior to the means by which cycling upon bound roots is prevented in the 'serial theories' Benua rejects. TCT accounts for the lack of cyclic identity effects caused by bound roots by the fact that bound roots never occur as outputs:

> Since a word built from a bound root has no output base (**electr, *ceive, *peach*), it can never show misapplication or "cyclic" effects (Benua, 1997: 5).

Benua employs a morphophonological constraint, BOUNDROOT, which prevents bound roots from occurring as outputs in their unaffixed forms under the ranking BOUNDROOT >> IO-MAX (Benua, 1997: 201). Benua's BOUNDROOT constraint ensures that an unaffixed bound root is never an output of the grammar. Instead, all unaffixed bound roots are subject to the 'null parse', e.g. *matern-* → ∅. Clearly, Benua's analysis is nothing more than stipulative: the only reason OO-correspondence constraints only apply between free words is because bound roots are stipulated to not be grammatical outputs. By virtue of its stipulation, TCT actually resembles the serial models it is supposed to improve upon. Indeed, Benua's analysis fares worse: because BOUNDROOT, like any OT constraint, is violable and rerankable between languages, Benua's proposal predicts that there should be languages where bound roots do occur as outputs.

OO-correspondence also faces a problem to which a stratal analysis is not susceptible – the fact that stems never surface alone. As noted by Orgun (1996a: 79-84, 205; 1996b: §5.2), there are cases where bound stems trigger paradigmatic identity effects. OO-correspondence cannot handle such cases – the theory predicts that paradigmatic identity can only be triggered by words.

7.5 Conclusion

In this chapter, the two means by which paradigmatic opacity, of which English stress preservation is an example, may be handled in Optimality Theory have been compared: TCT (Benua, 1997), and Stratal OT. This comparison has shown that Stratal OT is preferable to TCT in two main ways:

(i) Serialism is intrinsic in Stratal OT, but must be stipulated upon the null hypothesis of OO-correspondence in TCT.

(ii) Stratal OT relies upon just one single mechanism for modelling paradigmatic and non-paradigmatic opacity – strata – whereas TCT is limited to just the paradigmatic opacity of which stress preservation is an example.

In addition, criticisms levelled against Stratal OT (or similarly overtly serial models) by Benua proved to be false, and, in some cases, bounced back onto TCT:

(i) Constraint reranking between strata in Stratal OT falsely predicts that markedness constraints can rerank, cp. TCT.

- Reranking of markedness constraints between strata is in fact necessary – TCT is incorrect.

(ii) It is a bad thing that, in Stratal OT, paradigmatic opacity is a side-effect of the stratal organisation of the grammar and IO-faithfulness.

- In the stratal approach, only one type of faithfulness constraint is needed – IO-correspondence – and no additional and excessively powerful OO-correspondence constraints are required.
- Strata are motivated for reasons other than enforcing phonological paradigmatic opacity. Meta-rankings of OO-correspondence constraints (OO_1, OO_2) are not.

(iii) [mn]-simplification is incompatible with a stratal model.

- Bermúdez-Otero & McMahon's (2006) version of Stratal OT can handle [mn]-simplification.

(iv) OO-correspondence captures the observation that cyclic effects do not occur between bound roots, whereas in serial models this is a stipulation.

- This is a stipulation in OO-correspondence, just as it is in serial models.

In chapter 8, one of Benua's criticisms of stratal models – the use of both stratum-internal cyclicity and strata to model paradigmatic opacity – will be answered at length. This analysis will show that stratum-internal noncyclicity is vital if the true empirical nature of weak stress preservation is to be captured.

Chapter eight: weak stress preservation in Stratal Optimality Theory

8.0 Introduction

In this chapter, it is shown how the model of Stratal OT adopted in this thesis – that proposed in Bermúdez-Otero & McMahon (2006) and Bermúdez-Otero (in preparation) – handles weak stress preservation. In §8.1 and §8.2, this model's proposal of stem-level 'fake cyclicity' for the handling of weak stress preservation is outlined. Most importantly, it is shown how the notion of the fake cyclicity is supported by the frequency analyses of chapter 6. The fake cyclicity analysis is shown to be not only compatible with the frequency results, but necessary to handle them: some problems with models that retain the cycle in their handling of weak stress preservation, notably Marvin's (2002) model of Distributed Morphology, are noted in §8.3. In §8.4, it is shown that fake cyclicity can handle another process which has been argued to absolutely require the stratum-internal cycle – Trisyllabic Shortening. The chapter concludes in §8.5.

8.1 Fake cyclicity

In Lexical Phonology and Morphology ('LPM'), the highest morphophonological stratum is typically argued to be internally cyclic: every suffix that is added to a Stem (Root in Giegerich, 1994, 1999) triggers a pass through the stratum's phonology, so that phonology on the highest stratum is recursive. This stratum-internal cyclicity has been a feature of most models of LPM, with the notable exception of Borowsky (1993) and Giegerich's base-driven model of LPM (on the latter, see §1.2.1.2). Stratum-internal cyclicity is also retained in Kiparsky's (1998a, 2000, 2003a) model of Stratal OT.

Stratal OT as formulated in Bermúdez-Otero & McMahon (2006) and Bermúdez-Otero (in preparation) proposes a radical departure from the hypothesis of stratum-internal cyclicity. In this version of Stratal OT, the highest morphophonological stratum – the Stem level – is **internally noncyclic**. The phonology of each stratum consists of a single phonological cycle (in OT, a single pass through Gen and Eval) so that, in this respect, the operation of the stem-level phonology is identical to that of the word-level phonology.

By making stratum one internally noncyclic, Stratal OT resolves a weakness present in previous stratal models. In LPM, it was assumed that strata were internally cyclic unless there was good reason to think otherwise, or it was simply stipulated otherwise (Giegerich, 1994: 51-2, and references therein). As noted in §7.4.2, this characteristic of stratal models is criticised by Benua (1997). Benua correctly points out that, by adopting stratum-internal cyclicity, stratal models have two mechanisms for creating paradigmatic opacity: both inter- and intra-stratal cyclicity. The model of Stratal OT proposed here has the potential to resolve this lack of formal restrictiveness if, as it claims, there is no requirement for stem-level stratum-internal cyclicity: the interleaving of strata alone will account for all instances of phonological opacity, and so the model is maximally economical in terms of the devices it uses to handle paradigmatic opacity. (We also saw in §7.4.3.1 that the rejection of stratum-internal cyclicity solves the domain mismatch of [mn]-simplification.)

The rejection of stratum-internal cyclicity in Stratal OT is clearly desirable in order to defend it against Benua's criticism; however, we have to ask whether it is empirically plausible. There does appear to be at least one empirical phenomenon where stem-level-internal cyclicity is unequivocally necessary – weak stress preservation. Without a stem-level internal cycle, Stratal OT appears to have no mechanism to deal with weak stress preservation: the stratum-internal cycle is the means by which stress can be assigned to embedded stem-level forms like *original* in advance of stress assignment to embedding stem-level formations like *originality*, thus enabling preservation from *oríginal* to *orìginálity*. The alternative to stratum-internal cyclicity for modelling weak stress preservation on the Stem level – the introduction of OO-correspondence constraints onto this stratum – is clearly an undesirable solution to the problem: Stratal OT seeks to do away with such theories of 'horizontal' correspondence (chapter 7).

It will now be shown that the apparent paradox of preserving stress on an internally noncyclic stratum does not exist if we think of stem-level morphophonological relationships in a different way. The traditional cycle has the potential to function as an 'on-line' process: given the phonological domain structure of any complex stem, the cycle could derive that stem's stress pattern from scratch, at that moment. But, as we saw in chapter 6, there is good reason to think that the complex stems where weak stress preservation occurs are stored in the permanent lexicon:[1] the frequency results suggested a need for the option to access these morphologically complex forms directly, without reference to their composite parts, in order to account for stress preservation failure.

If the complex stems where weak stress preservation occurs are stored in the permanent lexicon, and these stems are permitted to have predictable phonological information like stress stored with them too (as argued shortly), then the cycle's ability to function as an on-line process will be superfluous: stress will be part of a word's lexical entry (Bermúdez-Otero & McMahon, 2006). As is now shown, the consequence of this lexical storage is that no stratum-internal cycle is required to capture the weak stress preservation generalisation on the Stem level: all stem-level outputs can be derived in a single pass through the stratum's phonology, and the apparent existence of stratum-internal cyclicity is fake.

In §8.1.1, the operation of fake cyclicity is exemplified. In §8.1.2, the fake cyclicity analysis of weak stress preservation is defended.

8.1.1 The fake cyclicity analysis of stem-level stress preservation

Here, the operation of fake cyclicity as presented in Bermúdez-Otero & McMahon (2006: §3.4) for weak stress preservation in *phenómenon* → *phenòmenólogy* is exemplified.

Bermúdez-Otero & McMahon (2006) propose that all stem-level outputs, including predictably complex stems, are stored in the permanent lexicon as lexical entries. Included in a lexical entry is the metrical structure assigned to the stem-level output; this argument is assumed for now, and defended in §8.1.2.

Because metrical structure is in the lexical entries of stem-level outputs, even preserving stress can be derived in a single cycle. In a traditional cyclic analysis, two

[1] Following Bermúdez-Otero (in preparation), I use the term 'permanent lexicon' to refer to all the morphological items a speaker memorises as part of his/her long-term memory. The term 'lexicon' in this sense is different from that used with respect to LPM, where the lexicon can house on-line processes as well.

passes through the phonology would be needed to achieve the preserving stress pattern of a complex stem like *Elìzabéthan*: stress would first be assigned to *Elizabeth*, then to *Elizabethan*, as in (1a), below. However, if lexical entries contain metrical structure, the first cycle which assigns stress to *Elizabeth* is avoided, as in (1b):

(1) Weak stress preservation: stratum-internal cyclicity vs. fake cyclicity

MAX-FootHead = the output correspondent of an input foot must be a foot head.

ALIGN(ω, L; Σ, L) = the left edge of the prosodic word aligns with a foot

NONFIN = the final syllable in the prosodic word must not be a member of a foot

(a) Stratum-internal cycle: *Elizabethan* = 2 phonological domains

Phonology: 1st cycle[2]

Input: /*Elizabeth*/	MAX-FootHead	ALIGN(ω, L; Σ, L)	NONFIN
☞a. E(lí.za)beth		*	
b. Eli(zá.beth)		**!	*

Phonology: 2nd cycle

Input: /*E(lì.za)beth-an*/	MAX-FootHead	ALIGN(ω, L; Σ, L)	NONFIN
a. (E.li)za(bé)than	*!		
☞b. E(lì.za)(bé)than		*	

(b) Fake cyclicity analysis: *Elizabethan* = 1 phonological domain

Phonology: only cycle

Input: /*E(lí.za)beth-an*/	MAX-FootHead	ALIGN(ω, L; Σ, L)	NONFIN
a. (È.li)za(bé)than	*!		
☞b. E(lì.za)(bé)than		*	

[2] I assume that a constraint not shown – ALIGN-R 'main stress is on the rightmost syllable of the word' – rules out the candidate ****Élizabeth***.

In (1b), cp. (1a), the preserving stress of *Elìzabéthan* is derived in a single cycle. The noun stem *E(líza)beth*$_N$ is listed in the speaker's permanent lexicon as a lexical entry. When the speaker has need to use *Elizabethan* for the first time, -*an* will be added to the stored noun stem *E(líza)beth*$_N$. A previous cycle in which the stress of *E(líza)beth*$_N$ is derived from an unmetrified underlying representation, /*Elizabeth*/, need not occur as part of this derivation, as the stress of *E(líza)beth*$_N$ is already known. In this phonological evaluation, the ranking of MAX-FootHead above ALIGN(ω, L; Σ, L) ensures that the output *E(lìza)(bé)than* wins: *E(lìza)(bé)than* is faithful to the foot-head marked on the input noun stem *E(líza)beth*$_N$. Weak stress preservation is achieved in a single pass through the stratum one phonology – no stratum-internal cycle is required.

Remember, however, the starting argument that all stem-level outputs are listed as lexical entries in the speaker's permanent lexicon. This means, therefore, that *E(lìza)(bé)than*$_A$ will itself be listed as a lexical entry in the speakers permanent lexicon. The implication of this is that the phonological derivation of *E(lìza)(bé)than*$_A$ from *E(líza)beth*$_N$ + -*an* in (1b) is a 'redundancy rule' (Jackendoff, 1975):[3] the speaker does not need to perform the derivation in (1b) when he or she wants to use *Elizabethan* in future, but can go straight to the stored stem-level output *E(lìza)(bé)than*$_A$, which can then proceed directly onto the Word level (where it receives the necessary inflection to allow it to enter the Phrase level).[4]

The persistence of the redundancy rule, or, rather, the 'redundancy constraint ranking' (1b), that relates *E(líza)beth*$_N$ to *E(lìza)(bé)than*$_A$ is vital: it ensures that the stress preservation generalisation need not be sacrificed along with the rejection of stratum-internal cyclicity. The speaker will not perform the phonological computation from the underlying representation each time he wishes to use a stem-level output, but will, for example, be aware that the underlying representation of *Elìzabéthan* is not simply morphologically unanalysed *E(lìza)(bé)than*$_A$, but [[*E(líza)beth*]$_N$ -*an*]$_A$. This knowledge means that the stress preservation

[3] Somewhat presciently with respect to this proposal, Hayes (1981: 145) proposes that stress may be stored in the lexicon, so that his stress rules "might be regarded in a sense as lexical redundancy rules, despite their rather derivational appearance".
[4] Ricardo Bermúdez-Otero (personal communication) points out that stratum one forms like *E(lìza)(bé)than*$_A$ are morphologically and phonologically already stems. Given this, there is no reason to assume that they will pass vacuously through the phonology every time they are used: as stems (c.p. roots), they are eligible to proceed directly onto the Word level.

generalisation is part of the speaker's continued linguistic competence, something very important indeed. As we saw in chapters 4 and 5, there is considerable empirical support for weak stress preservation; any model which did not have the stress preservation as part of a speaker's linguistic competence would, therefore, be incomplete and incorrect. A model which does not include the preservation generalisation as part of a speaker's linguistic competence would also make the wrong predictions regarding the effects of word frequency upon preservation. As noted in §1.1, Selkirk (1980: 597-8) proposes that the relationship between an embedding word and its base is purely historical – there is no synchronic stress preservation generalisation. In Selkirk's model, therefore, the likelihood of failure of stress preservation is predicted to depend only on the frequency of the derived word itself. This prediction is clearly at odds with the findings of chapter 6, where it was shown that the probability of preservation failure depends on the relative frequencies of the embedding word and the embedded word. There is also a strong psychological motivation for having the weak stress preservation generalisation as part of a speaker's linguistic competence. By maximising phonological similarity, preserved stress aids morphological transparency by assisting perception of the embedded word within the embedding word (Cutler, 1980, 1981; Kenstowicz, 1996). Literature on morphological processing notes that transparent phonology aids morphological decomposition (e.g. Fraunfelder & Schreuder, 1991: 173). The stress preservation generalisation may also aid the memorisation of the stratum one forms involved – see §8.1.2.1.

The stem-level constraint hierarchy is therefore anticipated to do very little on-line work: well-formed outputs are listed on the Stem level ready-to-go. Nevertheless, as described, the stem-level constraint hierarchy is vital for retaining the generalisation of weak stress preservation on stratum one. It is also worth noting that the persistence of the stem-level constraint hierarchy is vital in light of OT's principle of Richness of the Base. Richness of the Base requires that the constraint hierarchy alone predicts well-formed outputs, with no restriction upon the input. The stem-level constraint hierarchy will predict that, for any given input, the output will be a well-formed stem-level output of English: stress patterns like **ci(tý)* and **(cítro)nella* would never be well-formed stems in English so long as the stem-level constraint hierarchy is in place (Bermúdez-Otero & McMahon, 2006: §3.4). (For references and further discussion see Bermúdez-Otero (1999: 124, f.n. 47), and Bermúdez-Otero & McMahon (2006: §3.4).) Again, the listing of all stem-level outputs in the permanent lexicon does not obviate the function of the stem-level phonology.

8.1.2 Defence of the fake cyclicity analysis

The fake cyclicity analysis of weak stress preservation vitally requires that stem-level outputs are listed in their entirety, ready-to-go, including metrical structure.

As will be discussed further in §8.2, the proposal that the outputs of the highest morphophonological stratum are listed is not a new innovation by Stratal OT, but was already present in LPM. For example (see also §8.1.2.1), the model of LPM proposed by Kiparsky (1982) argued that every phonological output of the Stem level – including metrical structure (e.g. 1982: 50) – was listed as an 'identity rule', one consequence of which was that stratum-one stress preservation was ensured (§8.2.1, below). On the morphological side, Giegerich (1999) conceives of the morphology of the highest morphophonological stratum (in his model, the Root level) as a network of relationships between the listed morphological outputs of the stratum. Nevertheless, both Kiparsky and Giegerich retain the stratum-internal cycle on stratum one. Stratal OT's particular innovation is using the listing of stem-level outputs to reject stratum-internal cyclicity.

Although the listing of stem-level outputs is therefore not a new innovation by Stratal OT, it is still worth taking some time to defend Bermúdez-Otero's proposal further – fake cyclicity relies so crucially upon it. In §8.1.2.1, I give theory-external psycholinguistic evidence which could support the listing of stem-level outputs. In §8.1.2.2, a particular argument for the listing of stem-level metrical structure from Bermúdez-Otero & McMahon (2006), which comes from the behaviour of exceptional #LLL monomorphemic words like *Epàminóndas*, is presented.

8.1.2.1 The psycholinguistic argument

The early generative analysis of SPE placed a premium upon redundancy-free storage and formal economy (Giegerich, 1999: 156-8). Early Generative Grammar assumed that storage is as minimal as possible, and that computation costs little; additionally, under the principle of Occam 's Razor, there was no need for redundant storage in addition to computational mechanisms (ibid.). These assumptions have been relaxed in more recent Generative Phonology: for example, in LPM, Kiparsky (1982), Mohanan (1986: §2.6), Borowsky (1993) and Giegerich (1999: 158) all propose that the outputs of the highest stratum's morphological and phonological rules are stored in a speaker's memory as lexical representations. Importantly, as far as our enterprise here is concerned, these lexical representations contain predictable phonological features like stress. The argument is clearly formally uneconomical: any stratum one output can be computed by the stratum one morphology and phonology, but these outputs are also redundantly listed in the permanent lexicon.

It is precisely the formally uneconomical scenario that appears to resemble real-world facts. The highly successful dual-route theory of lexical access discussed in chapter 6 (evidence illustrating this model's success was cited in §6.2.1) proposes that a speaker can access any word he has heard before either via decomposition, or by directly recalling the memory of the entire word: lexical access is fail-safe, not economical. Similarly, there appears to be no such premium upon human memory that would prevent speakers from storing predictably complex words (Derwing, 1990; Burzio, 1996: 124; Pinker, 1999: 152); in fact, there is experimental evidence which indicates that speakers have memory traces of regular complex words (Pinker, 1999: 153). As noted by Mohanan (1986), the presence of redundancy rules in addition to stored lexical entries is likely to be a useful, rather than cumbersome, part of linguistic competence: "[a] list of words related through lexical redundancy rules is easier to store than a list of unrelated words" (Mohanan, 1986: 53). This argument is supported by the insights of connectionist pattern associator models of memory (Pinker, 1999: 132).

Without a premium upon memory, but with an emphasis upon efficient retrieval, there seems to be no reason why predictable phonological features like stress cannot be memorised, along with their predictably complex words. In the context of a more general defence of the view that redundant information (semantic, syntactic and phonological) is stored in the permanent lexicon, Giegerich (1999) notes psycholinguistic evidence which supports the storage of predictable metrical structure:

> [I]n the tip-of-the tongue phenomenon (Brown and McNeill 1966), for example, speakers may recall the number of syllables, or the stress pattern, or other possibly non-underlying characteristics of a lexical item while experiencing an inability to find the lexical item itself [...] Such results and others (Linell 1979) clearly suggest that speakers memorise complete words (rather than morphemes), in a form that closely resembles classical phonemic representations (rather than more abstract underlying representations), enriched by relevant suprasegmental structure (Giegerich, 1999: 157).

Similar support from the tip-of-the-tongue phenomenon is given by James (1890), Kozlowski (1977), Rubin (1975) and Elbers (1985) (all cited in Brown, 1991). In sum, from a psycholinguistic perspective, it is highly plausible that stem-level outputs are listed in their entirety, including their metrical structure.

Crucially, it is proposed here that the wholesale listing of stem-level outputs is not just psycholinguistically plausible, but psycholinguistically necessary.[5] Following from Hay's (2003) analysis, itself based upon a dual-route model of lexical access, it was proposed in chapter 6 that complex words which are candidates for preservation are memorised by the speaker once coined or heard for the first time. It is this storage of complex words which provides the mechanism by which weak stress preservation can fail. As was argued in chapter 6, whole-word access of stored embedding words has the potential to weaken the relationship between embedding and embedded words like *miscègenátion* and *miscégenate*, or *antìcipátion* and *antícipate*, creating the opportunity for stress preservation failure: ***mì****scegenátion* and ***àn****ticipátion*. Without this whole-word storage, preservation failure as a result of frequency effects cannot be accounted for. With respect to more general properties of phonology and semantics, Mohanan (1986) makes the same argument for the necessity of listing some derived forms; notably, Mohanan's argument is made with reference to his stratal model of LPM:

> Once a derived word becomes part of the word list [=permanent lexicon], it is in principle allowed to drift away from the meaning and pronunciation

[5] This argument is also made in Collie (2007).

> predicted by the principles of the grammar, and develop partial of full opacities of various kinds. [...] **In a theory that does not allow a subset of derived words to be listed in the lexicon, these phenomena do not make sense** (Mohanan, 1986: 54; boldface SC).

The reader may quite rightly wonder why memorisation should be limited to stratum-one outputs. In light of the dual-route model of lexical access which is advocated here, it would be odd if speakers did not also have memories of stratum two forms to which they had been exposed, e.g. *hopelessness*, *nationalism*. Similarly, speakers must have memories of syntactic idioms, e.g. *to pull someone's leg* ('to tease someone'). Following Bermúdez-Otero (2007: §26-§28), I propose that that different strata are subject to different types of listing. Stratum one, the home of fake cyclicity, is subject to **nonanalytic listing**: outputs are stored in a morphologically unanalysed form, along with predictable phonological structure like syllabification and stress. Outputs of lower levels may also be stored, but this storage will be **analytic listing**: the stored outputs will be morphologically or syntactically analysed, and may not necessarily be specified for predictable phonological characteristics. The necessity of analytic listing is shown by syntactic idioms. While there is no doubt that the speaker must memorise the meaning of *to pull someone's leg*, the internal syntactic structure of this idiom must be available to the syntax if constructions like *Whose leg did he pull?* are to be accounted for in addition to constructions like *He pulled her leg*. As noted by Bermúdez-Otero (2007: §26-§28), only nonanalytic listing can produce fake cyclic effects; fake cyclic effects will not, therefore, occur on the Word and Phrase levels even if some of their outputs are listed.

8.1.2.2 The *Epàminóndas* argument

There is also a specific case from our data to think that the fake cyclicity analysis is appropriate. This comes from the exceptional location of secondary stress in monomorphemic words like *Epàminóndas*.

Bermúdez-Otero & McMahon (2006: §3.4) propose that the exceptional position of secondary stress in monomorphemic #LLL words like *Ep**à**minóndas* (cp. ***à**bracadábra*) is due the presence of "underlyingly specified foot heads" which override the preference for the prosodic word to begin with a foot. In §5.2.2, we tried to find a phonological generalisation that predicted these apparent exceptions, examining both the incidence of onset-less syllables word-initially and any potential role of the height of the vowel in the word-initial syllable. While both hypotheses seemed to have some potential, both were falsified by some examples and indicated that a degree of lexical specification was necessary. I therefore follow Bermúdez-Otero & McMahon in assuming the specification of an underlying foot-head as a solution to these exceptional stress patterns in #LLL monomorphemic words.

What is interesting about exceptions like *Epàminóndas* is that exceptional, contrastive stress – *Epàminóndas* – and preserved stress – *phenòmenólogy* – in #LLL words occurs in exactly the same place. The connection between lexically contrastive and 'cyclic' phonological properties was noted in §7.4.1.1. This principle, known as Chung's Generalisation, is repeated in (2):

(2) Chung's Generalisation

If a stem-level phonological generalization displays cyclic misapplication, then it also has lexical exceptions.

As shown by Kiparsky (2007a), Stratal OT can happily capture the connection between preserving and contrastive stress. The connection is predicted by Stratal OT, as both exceptional and preserving stress must be handled by IO-faithfulness:

(3) Exceptional and preserving stress in Stratal OT (from chapter 7, (13))

Input: /*phe(nó.me)non-ology*/	IDENT-IO (stress)	ALIGN-L
☞a. phe(nò.me)(nó.lo)gy		*
b. (phè.no)me(nó.lo)gy	*!	
Input: /*E(pá.mi)nondas*/		
☞ a. E(pà.mi)(non)das		*
b. (È.pa)mi(nón)das	*!	

Bermúdez-Otero & McMahon (2006) extend the logic of this parallel between contrastive and preserving stress to argue for fake cyclicity. In the fake cyclicity analysis of weak stress preservation, it is assumed that underlying foot-heads can override the preference for a word-initial foot in a pre-tonic #LLL sequence. As (3) shows, this hypothesis also makes the correct predictions for monomorphemic words which are not candidates for cyclic preservation: an underlyingly specified foot-head can override the desire for a word-initial foot in *Epàminóndas* (Bermúdez-Otero & McMahon, 2006: 400). If monomorphemic words did not bear out this prediction, then we would have a serious problem (Bermúdez-Otero, 2007b: 12): under Richness of the Base, any underlying foot-heads in monomorphemic words should also be able to block the *àbracadábra* stress pattern.

8.2 Blocking and fake cyclicity

So far, we have established that the storage of stratum one outputs as lexical entries is vital to the analysis of weak stress preservation under the fake cyclicity analysis. Intriguingly, it was established in LPM that the storage of stratum one forms as lexical entries was vital to the modelling of morphosemantic blocking under the Blocking Effect (Kiparsky, 1982; Giegerich, 1999, 2001). In this section, parallels are drawn between morphosemantic blocking and the fake cyclicity analysis of weak stress preservation. This exploration indicates that the reinforcement of weak stress preservation on stratum one should also be conceived of as characteristic of the Blocking Effect that is already independently established on this stratum.

8.2.1 The Blocking Effect

Blocking, in its most general conception, is described as by Aronoff (1976: 43) as “the non-occurrence of one form due to the simple existence of another”. Most analyses tend to go a little further than this, assuming that the form which does the blocking is more specialised in some way – by virtue of being listed in the lexicon, idiosyncratic and/or less morphologically complex – than the form being blocked.

The generative literature on the topic of the Blocking Effect focuses upon the interaction of morphology and semantics, in particular, the interaction between regular and irregular morphology (e.g. Aronoff & Anshen, 1998: 239-240; Pinker, 1999; Giegerich, 2001). Blocking can account for gaps in what are otherwise (more) regular word formation processes. For example, while an adult speaker may understand outputs of regular word formation processes like *womans* and *gooses* as the plurals of *woman* and *goose*, he or she will regard the irregular plurals *women* and *geese* as more correct. Similarly, although an adult will understand what is meant by *furiosity* or *furiousness*, he or she will prefer *fury*. In each case, the speaker's knowledge of more irregular forms like *women*, *geese* and *fury* blocks him or her from using a more regular word formation process to produce *womans*, *gooses*, *furiosity* or *furiousness*.

Blocking is not a fail-safe phenomenon: sometimes speakers do use overregularised forms like *womens*. Giegerich (2001) shows that instances of genuine blocking failure can be explained if a form's ability to block another depends upon the storage of the blocking item in the permanent lexicon.[6] Evidence for blocking failure comes from the higher incidence of overregularisation errors in children as compared to adults. Children are far more likely to produce overregularised forms like *gooses* and *cooker* (the latter meaning 'someone who cooks') because they simply have not lived as long as an adult: the child will have fewer instances of exposure to potentially blocking forms like *geese* and *cook* (the latter meaning 'someone who cooks'), and, consequently, these forms are less entrenched in their memories. The dependence of blocking's success upon human memory means that it is an unavoidably probabilistic phenomenon (Aronoff & Anshen, 1998: 240), even in adults. Forms which have higher token frequencies are more likely to successfully block a default process than forms with low token frequencies.

In LPM, the psycholinguistic phenomenon of blocking has been crucially associated with stratum one. Both Kiparsky (1982) and Giegerich (1999, 2001) argue that blocking occurs because specifically the outputs of stratum one – the home of the irregular forms like *fury* and *geese* – are listed in a speaker's permanent lexicon. These lexical entries block later (more) regular word formation, e.g. stratum one *fury* blocks stratum two **furiousness*, or more complex stratum one **furiosity*.

---

[6] Arguments for putative 'type blocking' – where blocking is confined to specific derivational paths – do not rely upon the blocking item being listed in the permanent lexicon (Giegerich, 2001). Giegerich (2001) makes a convincing argument against type blocking, and argues that all instances of blocking are of the 'token'-type described here: a memorised form does the blocking.

8.2.2 Blocking, fake cyclicity and weak stress preservation

From the discussion in §8.2.1, it is clear that blocking is a process which vitally involves what are termed here 'lexical entries': lexical items, notably outputs of stratum one, listed in a speaker's permanent lexicon. It was also observed that blocking is a probabilistic phenomenon dependent upon frequency effects: the success of blocking depends upon how many times a speaker has been exposed to the blocking form. In light of the fake cyclicity analysis presented so far in this chapter, and the frequency results from chapter 6, weak stress preservation is also amenable to an analysis in terms of blocking.

Under the fake cyclicity analysis, all stem-level outputs are listed in a speaker's permanent lexicon, their stress patterns included.[7] A preserving form like *phenòmenólogy* is a lexical entry in its own right, and this can account for why, even in the absence of the stratum-internal cycle, we do not get non-preserving *phènomenólogy*: *phenòmenólogy*, by virtue of being listed in a speaker's permanent lexicon, blocks the existence of *phènomenólogy*.

Importantly, the absence of *phènomenólogy* cannot be explained by the absence of a word formation process which could create it, in the same way that the absence of the word *furiosity* in §8.2.1 could not be explained by the absence of a word formation process that could create it. *Phènomenólogy* is a possible stratum one output: the suffix *-ology* can attach to bound roots, e.g. *proctology* (Ricardo Bermúdez-Otero, personal communication), and it is also possible that a speaker will perceive a bound root *phenomen-*$_R$ by comparing words like *phenomenal* and *phenomenon* (ibid.). In the absence of a stratum-internal cycle, direct suffixation of the bound root *phenomen-*$_R$ with *-ology* is a possible output of stratum one word formation, and the stress pattern associated with this form would be *phènomenólogy* (bound roots cannot logically be stressed).

Where this blocking account of stem-level stress preservation becomes really interesting is in its ability to account for stress preservation failure. We noted in §8.2.1 that blocking is probabilistic: occasionally, **furiosity* or **furiousness* might

[7] The particular account of blocking and weak stress preservation given in this paragraph and the next is the result of personal communication with Ricardo Bermúdez-Otero and consultation of Bermúdez-Otero (2007b: §32). The analysis of blocking failure which follows is, however, primarily my development.

occur if the speaker fails to retrieve *fury* from his memory in time to block both of them. The frequency effects seen in chapter 6, in combination with the model of lexical access discussed there, mean that stress preservation failure can also be seen as an instance of blocking failure.

To recap (see §6.2.1 and §8.1), under the dual-route theory of lexical access adopted in this thesis, two methods of lexical access are proposed: the decomposed route, by which a complex word is accessed via its composite parts, and the whole-word route, when a complex word is accessed directly, without decomposition into its composite parts. As discussed in detail in chapter 6, Hay (2001, 2003) hypothesises that when complex words are more frequent than their immediately embedded words, they are prone to the whole-word route, and so 'forget' the semantics and phonology of their embedded words – the embedded word is not being referred to during the access of the embedding word.

In the context of stress preservation, repeated whole-word access of *miscegenation* is prone to weaken its redundancy rule relationship with *miscégenate*. If a speaker does not recall *miscégenate* in the lexical access of *miscegenation*, he is unlikely to perceive *miscegenation* as consisting of *miscégenate*$_{V}$+ *-ion*. Consequently, the speaker may still recognise *miscegenation* as complex in some way – namely, a bound root *miscegen-*$_{R}$ suffixed with the fused suffix *-ation* (as in *trepid*$_{R}$ + *-ation*) – but the underlying representation in this case would just be [[*miscegen*]$_{R}$ *-ation*]$_{N}$.[8] Without the stress of *miscégenate* present in the underlying representation of *miscegenation*, the stratum one phonology predicts the output *mìscegenátion*, not *miscègenátion*. It has been established that one of the key proposals of fake cyclicity is that the preserving output *miscègenátion* is stored; it is not the case, therefore, that blocking failure will immediately lead to stress preservation failure as in *miscegenation* (see also §6.3). However, blocking failure will make the form subject to the pressure of gradual lexical change, so that, in time, we may see a move towards the phonologically well-formed, non-preserving pattern.

The important implication of the probabilistic weakening of the redundancy rules between embedding and embedded words for our blocking analysis of stress preservation is this: by the weakening of the relationship with *miscégenate*, the

[8] On the change in the underlying representation, see §9.1.

lexical entry *miscègenátion* will no longer be present to block the existence of the possible stem-level output *mìscegenátion*. The similarity between this blocking account of stress preservation and the account of Synonymy Blocking is clear: just as the non-existence of a possible form like *furiosity* is probabilistically dependent upon *fury* being at a sufficiently high resting-activation level, and therefore token frequency, to block it, so is the non-existence of a possible form like *mìscegenátion* probabilistically dependent upon *miscégenate* being of a sufficiently high frequency to block it.

The similarity of the mechanisms which enforce Synonymy Blocking and stratum one stress preservation was already implicit in LPM, where both were ultimately attributable to the Elsewhere Condition (Kiparsky, 1982; Giegerich, 1999: 102-3):

(4) Elsewhere Condition (Giegerich, 2001)[9]

The processes A, B apply disjunctively iff:
(i) A is restricted to a single lexical item; B is not so restricted
(ii) A and B are rival processes such that *either*
(a) the input of A answers the structural description of B and the outputs of A and B are distinct; *or*
(b) the output of A is equivalent to that of B and the inputs of A and B are distinct

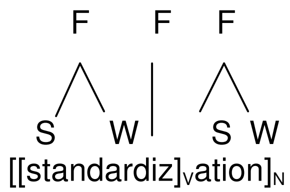
On the morphological side, the Elsewhere Condition predicted Synonymy Blocking. But, vitally, on the phonological side, the Elsewhere Condition manifested as Strict Cyclicity: structure changing rules apply only in derived environments, and only structure building rules can apply in nonderived environments. In the model of LPM proposed by Kiparsky (1982), the Elsewhere Condition, via Strict Cyclicity, ensured stress preservation (Kiparsky, 1982: 49-52), as is now shown.[10]

In Kiparsky (1982), the output of each stem-level phonological cycle, metrical structure included, was listed as a lexical entry: ‘identity rules’. These rules constitute the more specific ‘rule A’ of the Elsewhere Condition. The stress rules

[9] This version of the Elsewhere Condition is used because Giegerich (2001) shows it to predict both Synonymy Blocking and Strict Cyclicity correctly.
[10] I use Kiparsky here, rather than Giegerich (1999), as Giegerich’s base-driven model of LPM suffers serious problems when faced with stratum one stress preservation (§1.2.1.2).

which assign primary and secondary stress constitute the more general Rule B. Consider this example of preservation given by Kiparsky:

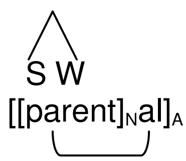
(5) Kiparsky (1982a: 50)



On the cycle in which *standardization* is derived, the grammar is faced with the footed lexical entry of *standardize*$_V$, and the unmetrified suffix -*ation* added on that cycle. Kiparsky's English Stress Rule can apply to -*ation* no problem: -*ation* corresponds to no stored lexical entry, and so there is no Rule A to block the standard application of stress rules: Rule B. This application of stress is structure building, not structure changing. However, when the secondary stress rule (iterative Strong Retraction) comes to apply to the left of the main stress in *standardization*, the situation is altogether different. The domain of the secondary stress rule's application is the lexical entry *standardize*$_V$, and to overwrite this pattern would be structure changing in a nonderived (as far as this cycle is concerned) environment (Kiparsky, 1982: 50). The lexical entry *standardize*$_V$ therefore blocks the application of secondary stress, giving the correct pattern *stàndardizátion*, rather than the incorrect pattern **standàrdizátion* Kiparsky (1982: 49) argues to be predicted by the secondary stress rule in the absence of metrical-structure preservation.

Strict Cyclicity does not always predict the preservation of metrical structure from previous cycles, as evident in *párent* → *paréntal*:

(6) Kiparsky (1982a: 51)



On the cycle in which *paréntal* is derived, the English Stress Rule considers the (indicated) string -*rental* (Kiparsky, 1982: 51). This string corresponds to no stored lexical entry, and so the English Stress Rule is free to apply: -*rental* is a derived environment, and stress rules can apply in structure changing mode. The metrical structure of *párent* is not preserved.

Kiparsky's (1982) analysis obviously makes some different assumptions to ours about stress assignment (notably, *standàrdizátion* would not be a default stress-pattern in our analysis, re: chapters 3 and 4), but, importantly, his model links the Strict Cyclicity which ensures stratum one stress preservation to listing, an argument Giegerich (1999) is explicit about (following Borowsky, 1993): "there is, in English at least, a correlation between listing and SCE [Strict Cyclicity Effect]" (Giegerich, 1999: 105). This link between Strict Cyclicity and listing is important because, as noted, Strict Cyclicity is the phonological manifestation of the Elsewhere Condition, and the Elsewhere Condition and listing are also responsible for the Blocking Effect.

Kiparsky's insight that stratum one stress preservation crucially relies upon the listing of all stratum one outputs as lexical entries is retained in Stratal OT, as exemplified so far in this chapter. There is, however, the obvious major difference that Strict Cyclicity itself – vital to ensure stratum one stress preservation in the LPM model – has no place in the model of Stratal OT proposed here. The Elsewhere Condition is not a formal condition on the phonological grammar in OT (Prince & Smolensky, 2004 [1993]: 68, 131-2) (see §1.4.1). Blocking, instead of Strict Cyclicity, enforces stress preservation in Stratal OT: this is intuitive, given that in LPM the two mechanisms were already recognised as being closely related, and as being reliant upon listed lexical entries. Vitally, this revision has the advantage that blocking correctly predicts the probabilistic nature of stratum one stress preservation.

8.3 Fake cyclicity versus the cycle

So far in this chapter we have seen that, by conceiving of stratum one stress preservation in terms of fake cyclicity, we can capture its probabilistic nature. We will now see precisely how the probabilistic nature of weak stress preservation causes problems for a recent analysis of weak stress preservation which uses the cycle. Marvin's (2002) analysis of cyclic stress preservation in the framework of Distributed Morphology is focused upon here but, as discussed at the close of §8.3.2, the criticisms made with respect to Marvin's model extend to other analyses of weak stress preservation which retain the phonological cycle's complex phonological domain structure.

8.3.1 Distributed Morphology and stress preservation

Marvin (2002) proposes analyses of left-edge and putative pre-tonic stress preservation in the framework of Distributed Morphology (Halle & Marantz, 1993; Marantz, 1997). Distributed Morphology ('DM') is an anti-lexicalist theory of morphology: syntax is used to construct sentences and words alike. Concomitant with this is that DM rejects the assumption of interactionism that is at the heart of Lexical Phonology and Morphology: as in SPE, word formation and phonological processes are not interleaved, but, rather, all syntactic concatenation applies first, before the phonological and semantic interpretation. This is diagrammatised in figure 1:

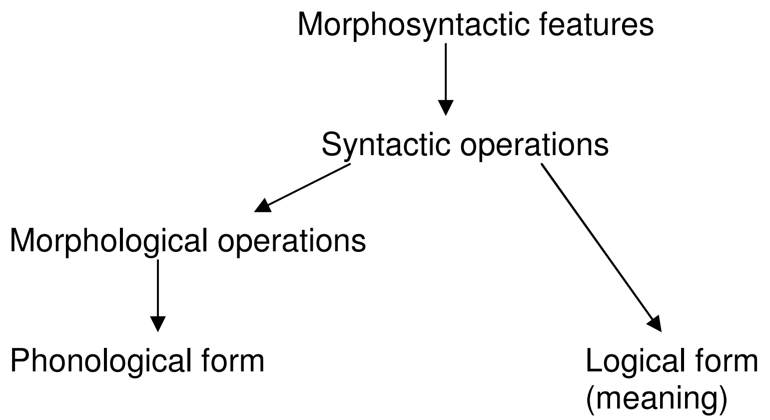


Figure 1: The grammar in Distributed Morphology
(simplified version of Harley & Noyer (1999: 3))

Marvin uses the process of 'Derivation by Phase' (Chomsky, 2001), proposed in the Minimalist Program, to model stress preservation in her model of DM. Derivation by Phase is analogous to the SPE transformational cycle. Summarising Marvin (2002: 18-19), sentences are built step-by-step, phonology taking place at certain points – 'phases' – in the syntactic tree: at each phase, the sentence is 'shipped' to 'Phonological Form' ('PF') and 'Logical Form' ('LF') ('spell-out').

Vitally, Marantz (2001) proposes that Derivation by Phase applies within words, as well as above the word level (Marvin, 2002: 17).

Cyclic stress assignment operates in Marvin's model as follows. Like Halle & Vergnaud (1987a), Marvin (2002) proposes that suffixes are diacritically marked as cyclic or noncyclic; addition of a cyclic suffix triggers a new spell-out phase which includes a pass through the cyclic stress rules. The Phase Impenetrability Condition (Marvin, 2002: 20) dictates that what is spelled out on a phase cannot be changed later in the derivation: phonology from previous cycles is preserved in the next cycle.[11] Marvin argues that stress assignment is subject to phase spell-out and the Phase Impenetrability Condition, and, in this way, brings about stress preservation: an embedding word has to retain the stress pattern of its embedded word, as the Phase Impenetrability Condition prevents the stress pattern of the embedded word from being overwritten.[12]

Marvin's analysis of stress preservation is a straightforward translation of the SPE transformational cycle into DM. Stress patterns must be strictly representative of the morphological structure of a word, and any apparent stress preservation failure cannot be any such thing: it must be accounted for by an alternative morphological analysis, and therefore phonological domain structure, which predicts this stress pattern. As is now shown in §8.3.2, this sort of cyclic analysis lacks explanatory adequacy when it comes to preservation failure.

8.3.2 Problems for Marvin's Distributed Morphology analysis

The potential for alternative morphological analysis as a way around stress preservation failure is particularly relevant here because our biggest examination of stress preservation failure – chapter 6 – dealt with *-ation* words, e.g. *anticipation*. Such words are prime targets for alternative morphological analyses: *anticipation* can be analysed not only as an affixations of a free word, and therefore a candidate

[11] The Phase Impenetrability Condition is therefore analogous to LPM's Strict Cyclicity.

[12] Marvin (2002: 57) can therefore only cope with the stress subordination seen in weak stress preservation by stipulating that the Phase Impenetrability Condition only holds at lines 0 and 1of the metrical grid. Otherwise, the stress shifting associated with weak stress preservation cannot be accounted for: the preserved stress would be predicted to be the most prominent in the embedding word. This is clearly an inferior analysis to an OT one presented here, where the preservation of stress at a lower level of prominence is grounded in the interaction of stress preservation constraints with markedness constraints.

for preservation – [[*antícipate*]$_V$ -*ion*]$_N$ – but also as an affixation of a bound root – [[*anticip*]$_R$ -*ation*]$_N$. This is because -*ation* can function as a fused suffix (e.g. [[*deforest*]$_V$ -*ation*]$_N$, **deforestate*) and can attach to bound roots (e.g. [[*dur*]$_R$ -*ation*]$_N$). Furthermore, bound roots cannot logically bear any stress. Therefore, if stress preservation from *miscégenate* or *antícipate* fails, and *mìscegenátion* or *ànticipátion* occurs instead, a cyclic analysis such as Marvin's may argue that these are affixations of bound roots.[13]

The problem with such a cyclic analysis of preservation failure is that it is simply descriptive: it gives no reason why we should get [[*anticip*]$_R$ -*ation*]$_N$ rather than [[*antícipate*]$_V$ -*ion*]$_N$. This is very much a problem: for the great majority of words ending in -*ation* where second syllable preservation is expected that were examined in chapters 4-6, stress is consistently in preserving position and does not fail, whereas the cyclic analysis predicts a failing option for all -*ation* words. In contrast, the fake cyclicity analysis can account for this preponderance of successful second syllable weak preservation: the fake cyclicity analysis can take frequency effects in to account. Under the fake cyclicity analysis, *anticipation* is predicted to have the morphological structure [[*anticip*]$_R$ -*ation*]$_N$ only when *antícipate* is not sufficiently frequent to be referred to in the lexical access of *anticipation*.

There is a second problem with the potential DM solution. As pointed out by Ricardo Bermúdez-Otero (personal communication), any traditional cyclic analysis must stipulate that -*ation* in [[*anticip*]$_R$ -*ation*]$_N$ is a single fused suffix, and not [[*anticip*]$_R$ -*ate*]$_V$ -*ion*]$_N$: if the latter was the morphological structure of *anticipation*, then a traditional cyclic analysis would predict that stress had to be preserved from *antícipate* which is created on the way. The fake cyclicity analysis suffers no such problem. Whether the structure of -*ation* is indeed fused -*ation*, or -*ate* + -*ion*, if the lexical entry *antícipate* is not sufficiently frequent, stress preservation will fail.

At this point, it should be made clear that Marvin's analysis is not alone in its failings: any analysis of weak stress preservation which, like the cycle, fails to recognise the probabilistic character of morphological decomposition, and therefore the gradient nature of morphological complexity, will struggle to deal with

[13] Marvin (2002: 69) takes precisely this approach with pre-tonic preservation failure. In chapter 2, it was argued that pre-tonic preservation is preservation of vowel quality, not stress. Marvin's example is therefore not used here.

probabilistic stress preservation failure in any principled manner. The criticism of Marvin (2002) therefore similarly applies to Benua (1997), presented in chapter 7, and, of course, the stratum-internal cycle proposed for stratum one in models of LPM.

There is a second, albeit more speculative, problem for cyclic analyses in addition to accounting for preservation failure: phonological characteristics which 'leapfrog' over cycles. Under the traditional cycle, we expect phonology to respect 'strict locality' (Benua, 1997): a complex form will only inherit the phonological features of its immediately embedded constituent. So, for example, *originality* will inherit the stress of *oríginal* to give *orìginálity*, not of more-deeply embedded *órigin* to give **òriginálity*. But, while strict locality is assumed in cyclic analyses, it does not automatically follow from theories of lexical access. In chapter 6, we discussed the way in which frequency affects lexical access: more frequent words have higher resting-activation levels in the permanent lexicon, and so are more easily accessed. As argued in §6.5.4.1, there seems to be no reason why this should be limited to embedding words and their immediately embedded words: if, for example, *tótal* is more frequent than *totálity*, *tótal* may influence the stress pattern of *totalitarian* more than *totálity* does. It may well be the case that what appears to be a non-preserving stress pattern under the strictly local analysis – *tòtalitárian* – is, in fact, stress preservation from the more-deeply embedded word *tótal*.

In §6.5.4.1, words were presented which are candidates for 'leap-frogging' stress preservation. These are repeated in (7) for Jones (2003), along with their CELEX token frequencies:

(7) Potential examples of leap-frogging preservation from Jones (2003)

	Embedding word (variable secondary stress)	Freq.	More-deeply embedded word	Freq.	Immediately embedded word	Freq.
1	ambassadorial	4	**ám**bassy (OED)	0	amb**á**ssador	279
2	antipathetic	5	**ànti**páthic (OED)	0	ant**í**pathy	23
3	coincidental	16	c**òin**cíde	189	co**ín**cidence	243
4	humanitarian	38 (adj) 7 (n)	h**ú**man	5113	hum**á**nity	334
5	iconoclastic	12	**í**con	62	ic**ó**noclast	5
6	Mephistophelean	2	Meph**í**sto	0	M**èphi**stópheles	3
7	Shakesperiana	0	Sh**áke**speare	357	Shakesp**é**rian	48
8	totalitarian	98	t**ó**tal	1997	tot**á**lity	58
9	triangularity	0	tr**í**angle	219	tri**án**gular	61
10	triangulation	0	tr**í**angle	219	tri**án**gulate *tri**án**gular?*	0 *61*
11	utilitarian	62 (adj) 2 (n)	**ú**tilise	131	ut**í**lity	154

While a much larger sample size would be required to obtain statistical evidence for leap-frogging preservation, for 8/11 of the embedding words in (7), the more-deeply embedded word is more frequent than the embedding word. It is therefore plausible that these more-deeply embedded words may be contributing to what appears to be preservation failure from the strictly local word.

Worryingly for advocates of the cycle, pretty concrete support for leap-frogging phonology comes from stratum one vowel shortening. In (8a, b), vowel shortening patterns expected under the cycle have been given (see §8.4 for a more in-depth phonological discussion). However, Bermúdez-Otero (2007b: 12) has also noted the leap-frogging pattern, (8c), in the speech of a former colleague, Dr. John Hutton, at the University of Manchester in the UK:

(8) Vowel shortening and the 'Manchester Paradigm'

(a)	cycle [aɪ]	cyclic [aɪ]	cyclicity [aɪ]
(b)	cycle [aɪ]	cyclic [ɪ]	cyclicity [ɪ]
(c)	cycle [aɪ]	cyclic [ɪ]	cyclicity [aɪ]

In both (8a) and (8b), the vowel in the initial syllable of *cyclicity* is the same as that in *cyclic*, with that in *cyclic* being determined by the variable metrification behaviour of -*ic* so as to cause variable trochaic shortening. However, (8c) – the 'Manchester Paradigm' – is altogether more problematic: *c*[aɪ]*clicity* has the same vowel as more-deeply embedded *c*[aɪ]*cle*, and has bypassed the intermediate, strictly local form *c*[ɪ]*clic*. The Manchester Paradigm is no problem for the fake cyclicity analysis of stratum one. Fake cyclicity acknowledges that speakers relate words probabilistically, as predicted by lexical access; it is therefore entirely plausible that a speaker may associate two words that are not strictly local. In contrast, (8c) is damning for a cyclic analysis where no direct relationship between *cycle* and *cyclicity* is predicted – (8c) is blatant evidence for the disrespect of the cycle's strict locality. In the conclusion of the thesis, I consider just how far this less restrictive notion of phonological identity should be extended.

In conclusion, the cycle's failure to recognise morphological complexity as gradient and probabilistic causes it major problems: stress preservation failure cannot be explained, and leap-frogging phonology cannot even be described.

8.3.2.1 An argument for OO-correspondence?

At this point, it is worth considering whether fake cyclicity is the only mechanism that can cope with leap-frogging phonology.

In chapter 7, we discussed the OT theory of OO-correspondence. There, we noted that OO-correspondence, in its null hypothesis, permits phonological identity relationships between words which are not in a strictly local relationship. OO-correspondence therefore has the potential to predict leap-frogging phonology.

There are two reasons why Stratal OT is preferable to OO-correspondence when it comes to the handling of stress preservation. First, the fake cyclicity analysis still captures the asymmetrical nature of weak preservation relationships: all morphophonological relationships are enforced through unidirectional IO-correspondence constraints. Second, the handling of phonological opacity by Stratal OT is better overall: Stratal OT can account for both paradigmatic and non-paradigmatic opacity, whereas OO-correspondence can account for paradigmatic opacity alone (§7.4.2).

8.4 Trisyllabic Shortening

In the model of LPM proposed by Giegerich (1999), Trisyllabic Shortening is a stratum one phonological process.[14] Like weak stress preservation, Trisyllabic Shortening seems to vitally require the stratum-internal phonological cycle. It is therefore very important that the fake cyclicity analysis proposed here is compatible with Trisyllabic Shortening.

Trisyllabic Shortening (SPE) is a process of vowel shortening that is argued to occur in primary-stressed syllables which are the head of a trisyllabic sequence, e.g. *ná****tional***, *sánity*, *sincérity*, *austérity*. TSS operates in conjunction with qualitative vowel-height alternations in accordance with Vowel Shift (e.g. McMahon, 1990), as in *n*/eː/*tion* → *n*[æ]*tional*. TSS results in vowels which are phonologically long in embedded words being short in the embedding words, as in *n*/eː/*tion* → *n*[ǽ]*tional*, *s*/eː/*ne* → *s*[ǽ]*nity*, *sinc*/iː/*re* → *sinc*[έ]*rity*, *aust*/iː/*re* → *aust*[έ]*rity*.

TSS is argued to be conditioned by morphological complexity. In the trisyllabic environment in nonderived words, TSS is argued not to occur: e.g. /éɪ/*braham*, /óː/*beron*, /óː/*maha*. TSS was therefore classed as a structure changing cyclic rule that was limited to derived environments by Strict Cyclicity in LPM (Kiparsky, 1982; McMahon, 1990; Giegerich, 1999). It is recognised in LPM that some derived words constitute lexical exceptions to TSS, e.g. *ob*/iː/*se* → *ob*[iː]*sity*, not **ob*[ɛ]*sity*. Arguments in favour of the Strict Cyclicity account of TSS regard the presence of these lexical exceptions as merely characteristic of lexical rules, although

[14] As in chapter 1, Giegerich's model of LPM is focused upon because it resolves many of the flaws of earlier models of LPM, and its base-driven structure is that adopted in Bermúdez-Otero's (in preparation) model of Stratal OT .

Szpyra (1989: 74-5) argues them to be altogether more troublesome for the LPM analysis.

The argument for TSS being a crucially cyclic phenomenon comes from paradigms like *n*/eː/*tion*, *n*[ǽ]*tional*, *n*[æ̀]*tionálity*. Giegerich (1999: 101), following Kiparsky (1982: 42), argues that the length and quality of the initial vowel of *n*[æ̀]*tionálity* can only be accounted for if this vowel is inherited from *n*[ǽ]*tional*: "no rule is available to shorten the first vowel of *nationality* itself" (see also Giegerich, 1994: 51). In words which are suprasegmentally similar to *nationality*, but where there is no embedded form which fits the conditions for TSS, e.g. *h*/aɪ/*pochondria*, there is no shortening of the vowel of the initial syllable. Similarly, while *hypocrisy* has a short initial-syllable vowel which it inherits from the TSS form *hypocrite*, the initial-syllable vowel is long in *hypothesis* where there is no embedded form to which TSS has applied (Kiparsky, 1982: 42; Giegerich, 1999: 101).

As will have been clear from the discussion of left-edge stress preservation earlier in this chapter, the fake cyclicity analysis does not prohibit morphophonological relationships between stratum one forms. It is therefore very much feasible that *nationality* could have inherited the quality of its initial-syllable vowel from *national*. However, as I show here, one possible implication of Bermúdez-Otero & McMahon's (2006: §4) Stratal OT analysis of shortening in the TSS environment is that fake cyclicity is not required to account for the short initial-syllable vowel of *nationality*: *n*[æ]*tionality* can be predicted without reference to *national*, by a phonological process which is not specific to derived words.

Like Kager (1993: 425) (see also Prince, 1990; Lee, 1996), shortening in the TSS environment in Bermúdez-Otero & McMahon (2006: §4) is subsumed under a more general process of trochaic shortening. Under trochaic shortening, the preference for the ideal moraic trochee, (ˈH) or (ˈLL),[15] results in vowel shortening. Trochaic shortening captures the TSS generalisation in the following way (following Kager, 1993: 425). In *nation*, the second syllable is extrametrical: *(ná)*<*tion*>. The

[15] On the iambic-trochaic law see, for example, Kager (1993), Hayes (1995) and Alber (1997). It should be noted that footings given elsewhere in this dissertation do not always adhere to a bimoraic minimum. For reasons discussed in §9.4, I do not propose a final analysis of English stress in this dissertation.

first syllable is therefore head of a monosyllabic foot, (ˈH), which is a perfect moraic trochee; hence, no shortening occurs. However, in *national*, the second syllable is no longer extrametrical – *(ná.tio)*<*nal*>. If the vowel of the first syllable is long, as in *n*/eː/*tion*, the word's disyllabic foot will be (ˈHL). (ˈHL) is disliked as a moraic trochee. In contrast, if the vowel of the first syllable is short – (*n*[ǽ].*tio)*<*nal*> – the foot will be (ˈLL): an ideal moraic trochee.

Here, building on the argument of Bermúdez-Otero & McMahon (2006), I propose that the trochaic shortening argument extends to *nationality*: as in *(ná.tio)*<*nal*>, the second syllable of *(nà.tio)(ná.li)*<*ty*> is not extrametrical, but will rather be the second member of a disyllabic trochee. The vowel of the initial syllable of *nàtionálity* is therefore independently predicted to be short by trochaic shortening, regardless of whether or not the vowel is inherited from *nátional*: having a short vowel in the initial syllable of *(n*[æ̀]*tio)(náli)ty* gives a perfect moraic trochee, (ˈLL). If the vowel of the initial syllable was long, as in *n*/eː/*tion*, a less undesirable trochee, (ˈHL), would occur.

Bermúdez-Otero & McMahon (2006: §4) propose that trochaic shortening can be captured on the Stem level in Stratal OT by ranking a constraint which bans (ˈHL) moraic trochaic feet – RHHRM ('Rhythmic Harmony'; Prince & Smolensky, 2004 [1993]: 70) – above the constraint which requires faithfulness to underlying morae, MAXμ. The ranking RHHRM>> MAXμ will affect all feet in stem-level forms, not simply primary stressed feet (cp. the traditional TSS rule). Therefore, as shown in (9), this ranking predicts the correct outcome for the primary stressed syllables of *nátion* and *nátional*, **and** the secondary stressed initial syllable of *nàtionálity*:

(9) Trochaic shortening on the Stem level[16]

RHHRM = *(ˈHL)

MAXμ = each mora present in the input has a correspondent in the output

[16] This analysis assumes that the constraint PARSE-σ will be high ranking enough to force the parsing of the second syllable in *national* and *nationality* (re: Lee, 1996: 99, f.n. 21).

Input: /*n*/eː/*tion*/	RhHrm	Maxμ
☞ a. (*n*/eː/)<*tion*>		
b. (*n*[æ])<*tion*>		*!
Input: /*n*/eː/*tion-al*/		
a. (n/eː/.*tio*)<*nal*>	*!	
☞ b. (*n*[æ].*tio*)<*nal*>		*
Input: /*n*/eː/*tional-ity/*		
a. (*n*/eː/.*tio*)(*na.li*)<*ty*>	*!	
☞b. (*n*[æ].*tio*)(*na.li*)<*ty*>		*

In this Stratal OT analysis, there is no way that shortening is restricted to derived words, in contrast to the LPM TSS analysis: RhHrm >> Max-μ is a phonological generalisation that applies over derived and non-derived stem-level forms alike. Bermúdez-Otero & McMahon (2006: §4) argue that the trochaic shortening generalisation is equally applicable to nonderived words, based on examples like *A(mé.ri)<ca>*, *de(vé.lo)<p>* and *de(cré.pi)<t>*. It seems quite reasonable that the generalisation RhHrm >> Maxμ applies to derived and nonderived forms alike: although Kiparsky (1982) argues that TSS should be restricted to derived environments using Strict Cyclicity, he nevertheless points out that the majority of nonderived words conform to the TSS generalisation (e.g. *alibi*, *sycamore*, *camera*, *pelican*, *enemy*, *Amazon*, *Pamela*, *calendar* (Kiparsky, 1982: 35)) rather than violate it (e.g. *ivory*, *stevedore*, *Oberon* (Kiparsky, 1982: 35)) (see Sainz, 1992: 178).

However, there are more significant problems for the trochaic shortening analysis. There are long vowels in nonderived *Oberon* and *hypochondria*, but the trochaic shortening analysis predicts that these vowels should be short. In contrast, the TSS analysis proposed in LPM correctly predicts these vowels to be long. Still, LPM's TSS analysis also suffers from problems – the short vowels in nonderived examples like *camera* and *pelican*. With so many underived words conforming to the TSS generalisation, it is unclear why the process of Trisyllabic Shortening should be restricted to derived words only; it could, as proposed by Bermúdez-Otero & McMahon, be a generalisation which holds over simple and derived words alike.

Exceptions are a problem likely to be faced by all attempts to handle shortening in the trisyllabic environment (e.g. Myers, 1987: 516; Yip, 1987: 469; Rubach, 1996: 212), and are not a disadvantage uniquely suffered by the trochaic shortening analysis proposed here for Stratal OT.

There is a problem shared by both the trochaic shortening and traditional LPM TSS analysis: both analyses predict that the long vowel in derived *obesity* should be short. An indication of how this exception could be handled in Stratal OT comes from Lee (1996), who adopts an OT approach to shortening in the trisyllabic environment which is very similar to the trochaic shortening analysis proposed here.[17] Lee suggests that derived exceptions like *obesity* could be solved by either "a lexically specific constraint ranking" (a cophonology analysis would be more amenable to Stratal OT) or a "lexically fixed metrical foot" (presumably *o(bé)si*<*ty*>) (Lee, 1996: 93).[18]

In conclusion, there is no indication that shortening in the TSS environment is likely to be more problematic for Stratal OT than LPM. The presence of fake cyclicity on the Stem level would allow the model of Stratal OT proposed here to capture the situation whereby the short vowel in the initial-syllable of *nationality* is inherited from *national*, argued by Giegerich (1999) to be incontrovertible support for TSS. However, it has been shown here that Bermúdez-Otero & McMahon's analysis of shortening in the TSS environment as part of a more general process of trochaic shortening, applicable to derived and non-derived words alike, means that the short vowel in the initial syllable of *nationality* is predicted independently of fake cyclicity, on purely phonological grounds.

Outside the TSS context, in the more general context of trochaic shortening, there is evidence which indicates that the fake cyclicity analysis is more appropriate than the stratum-internal cycle for the general handling of stratum one vowel shortening. This evidence comes from the Manchester Paradigm, already noted in §8.3.2. Trochaic shortening can be seen in stratum one *-ic* adjectives, e.g. *conic*

---

[17] Lee (1996: 92) has a positively-stated constraint LL, H, which states the desire for well-formed moraic trochees, performing a similar function to RHHRM. This constraint is ranked above a constraint MAX(Tense), the latter performing a similar function to MAXμ.

[18] Mits Ota (personal communication) has pointed out that it may be worth exploring whether relative frequency effects (re: chapter 6) could account for failures of TSS, e.g. high frequency of *obese* relative to *obesity*.

versus *cone*, *cyclonic* versus *cyclone*. Apparent evidence for the cyclicity of trochaic shortening can be seen in the word *cyclicity* itself. Trochaic shortening in *cyclic* is variable: pronunciations with both shortened and long vowels occur (Jones, 2003, s.v. 'cyclic').[19] A short vowel in the initial syllable of *cyclicity*, itself not in the trochaic shortening environment, can be attributed to trochaic shortening in embedded *cyclic* – this is shown in (10a). Similarly, if shortening does not occur in *cyclic*, the initial-syllable vowel of *cyclicity* should be long – (10b). However, the counter-cyclic Manchester Paradigm, (10c), also occurs:

(10) Stratum one trochaic shortening and the 'Manchester Paradigm'

(a)	*cycle* [aɪ]	*cyclic* [aɪ]	*cyclicity* [aɪ]
(b)	*cycle* [aɪ]	*cyclic* [ɪ]	*cyclicity* [ɪ]
(c)	*cycle* [aɪ]	*cyclic* [ɪ]	*cyclicity* [aɪ]

As noted in §8.3.2, (10c) simply cannot be handled by the cycle, but can be by fake cyclicity.

8.5 Conclusion

In this chapter, Bermúdez-Otero's notion of stratum one fake cyclicity has been introduced as the mechanism for enforcing weak stress preservation. It has been shown that fake cyclicity not only results in a more restrictive model of paradigmatic opacity in Stratal OT, but that it can also explain why weak stress preservation is probabilistically dependent upon token word frequency, as was shown to be the case in chapter 6. In contrast, it has been argued here that the cycle can sometimes describe, but can never explain, stress preservation failure. It has also been shown that the cycle also cannot cope with leap-frogging stress and vowel quality preservation, whereas fake cyclicity can.

Interestingly, although fake cyclicity is a new proposal, it has been shown here that the nonanalytic listing of stratum one outputs and blocking upon which it

[19] Under a trochaic shortening analysis, this variable behaviour can be attributed to variable extrametricality of the *-ic* suffix.

relies were already present in previous stratal models. Fake cyclicity is therefore a logical progression from previous work in the stratal tradition.

Chapter 9: the constraints for left-edge stress preservation

9.0 Introduction

In this chapter, the precise OT constraints which are required to handle left-edge stress preservation on stratum one are presented, and their ranking exemplified (§9.1-§9.3). The constraint ranking for #HHL and #HLL monomorphemic and bound-root base words is then addressed in §9.4.

9.1 Foot-head preservation and its failure

In this thesis, left-edge preservation in light-initial words, and heavy-initial words which are candidates for initial-syllable preservation, is argued to be an instance of foot-head preservation: only the first **or** the second syllable of the embedded word has been argued to be the head of a foot: *o(rí.gi)nal* → *o(rì.gi)(ná.li)ty*; *(cá.pi)tal* → (*cà.pi)ta(lístic)*; *(pér.me)able* → *(pèr.me)a(bí.li)ty*. The first and second syllables of these words contrast simply in terms of stressed versus unstressed; therefore, in order to ensure preservation, all the preservation constraint need do is preserve this binary distinction.

As noted in §8.1.1, we can ensure foot-head preservation simply by using the constraint MAX-FootHead ('MAX-FtHd') as proposed by Bermúdez-Otero & McMahon (2006) (§8.1.1): 'the output correspondent of an input foot must be a foot head'. As long as MAX-FtHd is ranked above all potentially conflicting markedness constraints – for example, with respect to second-syllable preservation, ALIGN(ω, L; Σ, L) – then we can be sure of foot-head stress preservation.

The interesting part of the analysis of foot-head preservation comes with stress preservation failure. There are three examples of variable stress preservation failure in incontrovertibly #LLL words from chapter 4 (other examples have been excluded where the weight of the first or second syllables was reasonably questionable): *miscégenate* → *miscègenátion~mìscegenátion*, *vatícinate* → *vàticinátion~vatìcinátion*, and *Bolívia* → *bòliviáno~bolìviáno*.[1] A completely standard, totally ranked OT constraint ranking can handle the failure of foot-head

[1] Note that, in chapter 4, examples like *boliviano* were argued to be complex in spite of their terminal element not being a recognised English suffix.

preservation in these words. In §8.1, it was proposed that successful foot-head preservation is dependent upon a foot head being present in the word's underlying representation. The necessary foot head will not be present if a speaker fails to associate a complex word with its embedded word as a result of frequency effects: e.g., if the speaker does not have the underlying representation of *miscegenation* as $[[$*mis(cége)nate*$]_{V}$ -*ion*$]_{N}$, but rather perceives it as consisting of a bound root plus suffix as in $[[$*miscegen*$]_{R}$ -*ation*$]_{N}$ (or even as monomorphemic). Under this argument, the variability of stress preservation seen in *miscègenátion~mìscegenátion* is therefore due to variation in the input, as shown in (1):

(1) Variable foot-head preservation

(a) Successful foot-head preservation

Input: */mis(cége)nate-ion/*	MAX-FootHead	ALIGN(ω, L; Σ, L)
☞a. mis(cège)(nátion)		*
b. (mìsce)ge(nátion)	*!	

(b) Failure of foot-head preservation

(i) Input: */miscegen-ation/* (or */miscegenation/*)	MAX-FootHead	ALIGN(ω, L; Σ, L)
a. mis(cège)(nátion)		*!
☞ b. (mìsce)ge(nátion)		
(ii) Input: */abracadabra/*		
a. a(bràca)(dábra)		*!
☞ b. (àbra)ca(dábra)		

In (1a), the speaker associates *miscegenation* with *miscégenate*, hence there is a foot head present in the input of *miscegenation*, and preservation occurs. In (1b, i), the speaker has failed to associate *miscegenation* and *miscégenate* as a result of the frequency effects discussed in chapter 6. Consequently, foot-head preservation fails because there is no foot head present in the input, paralleling monomorphemic *abracadabra* in (1b, ii).

A key question here is the implication of variation in the input for the content of the embedding word's underlying representation: we seem to be saying that, in order to account for intra-speaker variation, a foot head is both present and not present in a word's underlying representation, or that the speaker has two underlying representations for a word which they then choose between (the latter situation is explored by Anttila (2002b: 218)). I assume that this is not the case; rather, I propose that the underlying representation is dynamic, changing over time as a result of frequency effects. As discussed in §6.5.3, lexical access is probabilistic and morphological complexity gradient; the relationship between the embedding and embedded word is, therefore, constantly varying in its strength. As a result of this variable relationship, the presence of the embedded word and its stress in the underlying representation of the embedding word varies in degree, resulting in a variable input to the grammar, and so variable preservation.

In the rest of the chapter, I argue that not all of the variable success of left-edge stress preservation can be attributed to variation in the input. In the analysis of relative prominence preservation given below, it will be argued that the variable success of preservation can also be attributed to variation in the grammar – constraint ranking – itself (§9.3). The foot-head preservation just discussed will be argued not to be subject to such variation in the grammar. These two different types of variation – of the input and of the grammar – will be argued to be motivated by two separate factors: frequency, as already presented here, and a perceptual difference in the nature of foot-head and relative prominence preservation (cf. §6.5.1). First, in §9.2, I present the constraints needed for the analysis of relative prominence preservation.

9.2 Predicting relative prominence preservation

The analysis of relative prominence preservation is considerably more complex than that of foot-head preservation: it needs to take into account the preservation of the relative prominence of stresses on adjacent syllables – e.g. ***(sèn)(sá****tio)nal* → ***(sĕn)(sà****tio)(náli)ty*, not ***(sèn)(sá****tio)nal* → *****(sèn)(să****tio)(nálity)* – rather than just the presence or absence of stress.[2] It has also been noted earlier in the thesis (§6.5.1) that, once frequency effects are factored out, relative prominence preservation is more prone to failure than second-syllable foot-head preservation. Both observations

[2] The reader is reminded that, in chapters 3 and 4, relative prominence preservation was shown to be relevant to only second-syllable preservation in heavy-initial words.

are captured in the analysis of relative prominence preservation proposed here: revised and new constraints are proposed to capture the preservation of the relative prominences of foot heads, and variable constraint ranking, as proposed in Anttila's theory of Partially Ordered Grammar (Anttila, 1997, 2002b, 2007), is used to capture the intrinsic fallibility of relative prominence preservation.

9.2.1 A constraint for relative prominence preservation[3]

I am aware of no constraint proposed in OT to date which has been used to differentiate between the relative prominences of degrees of subsidiary stress under preservation. In optimality-theoretic analyses, the relative prominences of non-primary stresses are generally not distinguished between, following analyses in Metrical Phonology like Halle & Vergnaud (1987a).[4]

There is an existing stress preservation constraint proposed by Pater (1995, 2000) which can be extended to bring about relative prominence preservation. Pater's (1995, 2000) preservation constraint, STRESSIDENT,[5] is a gradient rather than categorical constraint: it can be violated to varying degrees. In Pater's analysis, this is used to differentiate between strong preservation, weak preservation and no preservation, as in (2):

(2) Gradient violations of STRESSIDENT in Pater (1995: 26, 2000: 257-8)

(a) No violations: The head of the prosodic word in the embedded word is the head of the prosodic word in the embedding word:

hópe → ***hópe****lessness* = primary → primary

---

[3] In terms of OT constraints, I follow Pater (2000) and propose that PARSE-σ forces the parsing of stray heavy initial syllables as monosyllabic feet to give *(sèn)(sà.tio)(náli)ty*, not *sen(satio)(nali)ty*: see the analysis of *Tòrbáy* in (18) of chapter 2. My proposal for the parsing of stray, heavy initial syllables that are not immediately pre-tonic, as in *(sèn)(sà.tio)(náli)ty*, requires PARSE-σ to rank above constraints like *CLASH (Pater (2000) only deals with PARSE-σ>>*CLASH-HEAD, as in the *Tòrbáy* example).

[4] See especially Pater (1995: f.n. 1) and Yamada's (2005) critique of Hammond's (1999a) OT account of English secondary stress. My own impression is also informed by consultation of the Rutgers Optimality Archive and of published OT work on English stress, as well as responses to a question posted on the LinguistList.

[5] The constraint is called STRESSIDENT in Pater (1995); it is renamed IDENT-STRESS in Pater (2000).

(b) 1 violation (*): The head of the prosodic word in the embedded word is only the head of a foot in the embedding word:

*cond**énse*** → *cònd**èns**átion* = primary → secondary

(c) 2 violations (**): The head of the prosodic word in the embedded word does not correspond with the head of a foot in the embedding word:

***á**tom* → *<u>a</u>tómic* = primary → nothing

Alber's (1998: 25) stress preservation constraint is similarly gradient in her account of German stress preservation, where, following Kager (2000) (in real time), the constraint goes by the name PK-MAX.

Here, it is proposed that an extra gradation be added to this existing gradient stress preservation constraint. This allows us to differentiate between whether the embedded foot head has secondary or tertiary relative prominence, as in (3).

(3) Gradient violations of STRESSIDENT: reformulated

(a) No violations: ***hópe*** → ***hópe**lessness*:

= primary → primary

(b) 1 violation (*): *sèns**á**tional* → *sĕns**à**tionálity*

= primary → secondary

(c) 2 violations (**): ***sènsá**tional* → ***sènsă**tionálity*

= primary → tertiary

(d) 3 violations (***): ***á**tom* → ***a**tómic*

= primary → nothing

This reformulation of STRESSIDENT – called STRESSIDENT(I) from now on – differs only minimally from Pater's and Alber's versions in that an extra level of sub-optimal preservation is possible: that where there is a foot on the second syllable, respecting foot-head preservation, but it is the least prominent foot in the whole prosodic word – (3c) ***sènsá**tional* → ***sènsă**tionálity*.[6]

Having established STRESSIDENT(I) as the relative prominence preservation constraint, we now need to establish which constraints are in conflict with it so as to cause preservation failure. With respect to the effects of stress shifting suffixation, (3b), and the failure to preserve even foot heads, (3d), the conflicting constraints are already known. The requirement to have primary stress near the right edge of the word – ALIGN-HEAD – outranks STRESSIDENT(I) in English to bring about stress shifting suffixation: *(sèn)(sátio)nal* → *(sĕn)(sàtio)(náli)ty*, as in (3b). Stress may also be completely deleted in the embedding word, (3d), due to the combined pressures of ALIGN-HEAD and the requirement for minimally binary feet (FTBIN): *átom* → *<u>a</u>tómic*. However, I am aware of no current constraint which could require the failure to preserve the relative prominence of two feet, while not threatening the correct positions of the feet themselves: (3c) *(sèn)(sátio)nal* → *(sèn)(sătio)(náli)ty*. The constraint which brings about the situation in (3c) must be a new innovation by my analysis of stress preservation, and is now presented in §9.2.2.

9.2.2 Markedness conflict for relative prominence preservation

In this section, a new constraint is established which expresses the preference for *(sèn)(sătio)(náli)ty* over *(sĕn)(sàtio)(náli)ty*. This constraint expresses a desire for failed relative prominence preservation in the presence of successful second-syllable foot-head preservation. Before coming up with a constraint that could desire such a situation, it is vital to consider its grounding.

[6] A lack of formal economy results from having both STRESSIDENT and MAX-FootHead in the grammar: the function of MAX-FootHead is a subset of the function of STRESSIDENT. However, this overlap in constraint function is apparently unavoidable. For example, one could avoid the overlap in the functions of STRESSIDENT and MAX-FootHead by proposing a novel categorical constraint to replace STRESSIDENT which is only concerned with preserving relative prominence contours. Under this totally new constraint, *(ànti)cipátion* would be no worse an example of preservation from *àntícipate* than *(àn)(tĭci)pátion*; expressing the preference for *(àn)(tĭci)pátion* over *(ànti)cipátion* would be the job of MAX-FootHead. However, all this revised analysis would do is transfer the overlap to #LLL words, which are the domain of foot-head preservation alone: *oríginal* → **òriginálity* is not only a failure of foot-head preservation, but also a failure to preserve the prominence contour of the first two syllables of *oríginal*. The use of MAX-FootHead and revised STRESSIDENT, proposed here, at least retains some continuity with previous work on stress preservation.

In §6.5.1, it was shown that, once the effects of frequency have been factored out, the incidence of relative prominence preservation failure is higher than the incidence of foot-head preservation failure. We considered two possible reasons why relative prominence preservation may be more prone to failure than simple foot-head preservation: phrase-level rhythmic readjustment (§6.5.4.2), and perceptual factors (§6.5.1). As we need further evidence to support the fact that our dictionary entries are contaminated by phrase-level characteristics, the perceptual argument will be focused upon here.

Ricardo Bermúdez-Otero (personal communication) has suggested that there may be a greater perceptual cost associated with the failure of foot-head preservation – stress preservation at all levels – than of relative prominence preservation alone. If foot-head preservation fails, then it is not only the stress contour of the embedded word which may be lost: the segmental content of the embedded word is at risk too, as any unstressed vowel is at risk of vowel reduction (*ir*[ǽ]*scible* → **ìr*/ə/*scibílity*).[7] The cost associated with the failure of relative prominence preservation alone, with foot-head preservation being successful, is altogether less severe: the stress contour of the embedded word will be obscured by being reversed – ***àntí****cipate* → ***àntĭ****cipátion* – but, as long as foot-head preservation is successful, the segmental content of the embedded word is at least guaranteed to be visible in the embedding word. This difference in perceptual cost could create a situation whereby foot-head preservation must absolutely be respected, but violation of just relative prominence preservation is more tolerable. Within OT, motivating the ranking of constraints from perceptual arguments has been proposed by Steriade (1997), Hayes (1998), and Côté (2000: 153-4) (indeed, see also the acoustic-perceptual analysis of vowel reduction discussed in chapter 2). Notably, Fleischhacker (2001) argues that the degree of perceptual similarity between output and input plays a vital role in OT phonology.[8] Similarly, in the context of stress preservation, it is proposed here that degrees of perceptual similarity between output and input drive the phonology.

---

[7] In accordance with the argument in chapter 2, it is not held that all unstressed vowels will automatically reduce; the absence of stress does, however, make reduction a possibility.
[8] Specifically, Fleischhacker considers to what degree the location of vowel epenthesis in an output affects its perceived similarity to the input in languages with anaptyxis-prothesis.

Given the perceptual argument for why relative prominence preservation may fail in the presence of successful foot-head preservation, the constraint which could trigger this situation must be established. The preference for ***àntĭ****cipátion* over ***ăntì****cipátion* is not expressible through standard alignment constraints in OT. Alignment constraints chiefly express the preference for having a particular prosodic category – foot, syllable, or prosodic word – aligned with the boundary of a morphosyntactic or prosodic category. Alignment constraints like ALIGN-HEAD and LEFTMOST do include a specification that it must be the strongest foot which is to be aligned, but I am aware of no equivalents which refer to the prominences of subsidiary stresses. A pre-OT device will therefore be addressed to find the markedness condition which prefers ***àntĭ****cipátion* to ***ăntì****cipátion*: Hayes' (1984) Quadrisyllabic Rule.

The Quadrisyllabic Rule, (4), was discussed in §3.2.1:

(4) The Quadrisyllabic Rule (Hayes, 1984: 46)

A grid is eurhythmic when it contains a row whose marks are spaced close to two four syllables apart.

The relevant 'level of scansion' for the Quadrisyllabic Rule is that immediately below the main stress level, the level below this being dealt with by the Disyllabic Rule (Hayes, 1984: 48; Gilbers & Schreuder, 2002: 11).

Although discussions of the Quadrisyllabic Rule chiefly consider phrasal contexts, Hayes (1984: 41-2) indicates that this principle of eurhythmy is also applicable in single words like *Ticonderoga* (see also Sainz (1992), discussed in §3.2.1). If the Quadrisyllabic Rule is applicable to single words, then it can express that non-prominence-preserving *sènsătionálity* is preferable to prominence-preserving *sĕnsàtionálity*, both nevertheless equally respecting foot-head preservation (figure 1):

(a)

```
           x
      x ---x
  x    x    x
  x   x  x x x x
```

sĕ̀nsàtionálity

eurhythmy ✗ (relatively)
prominence preservation ✓
foot-head preservation ✓

(b)

```
           x
  x ------- x
  x   x     x
  x   x x   x x x
```

sènsằtionálity

eurhythmy ✓ (relatively)
prominence preservation ✗
foot-head preservation ✓

Figure 1: The Quadrisyllabic Rule within words

In figure 1, (a) *sĕ̀nsàtionálity* is more eurhythmic than (b) *sènsằtionálity* because the grid marks of the former for the row immediately below the main stress level are three syllables apart, not two. Note this has nothing to do with the presence of foot heads at lower levels, but is limited to the level of scansion immediately below the main stress level.

Gilbers & Schreuder (2002: 11) note that the Quadrisyllabic Rule's output-orientation is prescient with respect to OT, in particular, OT's alignment constraints. Gilbers & Schreuder informally propose the Quadrisyllabic Rule as an optimality-theoretic constraint, 'QR'. Of particular interest is that Gilbers & Schreuder (2002: 11) show QR conflicting with the phonological faithfulness that preserves the stress contour of the input in phrasal metrics, as in (5):

(5) Interaction of faithfulness and the informal QR constraint[9]

Input: *Mississìppi Délta*	QR	Correspondence
Mìssissippi Délta		*
Mississìppi Délta	*!	

Here, it is proposed that the Quadrisyllabic Rule should be formalised as a gradient constraint in OT, with the following precise formulation:

[9] It is assumed that a constraint not shown in (5) demotes the stress of *Mississíppi* to non-primary prominence under embedding.

(6) QR (based on Gilbers & Schreuder, 2002: 11)

The second-most prominent foot in a word or phrase is ideally spaced four syllables away from the most prominent foot. Every syllable by which this condition is not met incurs one violation (*).

Constraints such as *LAPSE set the precedent for a constraint like QR: *LAPSE is another instance of a surface-orientated rhythmic principle noted in pre-OT metrical literature (by Selkirk, 1984) being formalised as a constraint in OT (by Elenbaas, 1999). Furthermore, there seems to be cross-linguistic motivation for the rhythmic principle captured by the Quadrisyllabic Rule which provides typological grounding for the QR constraint – see Hayes (1984: 58-9).

With the QR constraint in place, it can be seen how this constraint interacts with STRESSIDENT(I) to cause failure of relative prominence preservation, but not foot-head preservation. If the constraint QR is ranked above the relative prominence preservation constraint STRESSIDENT(I), then relative prominence preservation will fail, as shown in (7):

(7) Failure of relative prominence preservation caused by QR

Input: */(àm)(bássa)dor-ial/*	QR	STRESSIDENT(I)
☞a. (àm)(bǎssa)(dóri)al	*	**
b. (ǎm)(bàssa)(dóri)al	**!	*
c. (àmba)ssa(dóri)al	*	***!

Importantly, in (7), QR does not require the failure of foot-head preservation. As far as QR is concerned, both candidate a. *(àm)(bǎssa)(dóri)al* and c. *(àmba)ssa(dóri)al* are equally well-formed – both score one violation with this constraint. It is the greater violation of STRESSIDENT(I) by c. *(àmba)ssa(dóri)al* than by a. *(àm)(bǎssa)(dóri)al* which puts the former out of the running for the winning output.

Under the opposite constraint-ranking, STRESS-IDENT>>QR, relative prominence preservation is predicted, as shown in (8):

(8) Successful relative prominence preservation: STRESSIDENT(I)>>QR

Input: */(àm)(bássa)dor-ial/*	STRESSIDENT(I)	QR
a. (àm)(bǎssa)(dóri)al	**!	*
☞b. (ǎm)(bàssa)(dóri)al	*	**
c. (àmba)ssa(dóri)al	***!	*

In §9.3.2 of this chapter, it will be shown that both the rankings QR>>STRESSIDENT(I) and STRESSIDENT(I)>>QR play a role in the stratum one grammar. This is a different way of modelling variation to that presented in §9.1: here, variation occurs as a result of variation in the grammar rather than in the input. I argue this to occur because, as already noted, relative prominence preservation is variable in its success even when the effects of frequency – variation in the input – are factored out.

The QR constraint's predecessor, the Quadrisyllabic Rule, functioned as a readjustment rule upon pre-existing foot heads: stress can shift in *bàmbóo*, as in *bàmbŏo cúrtain* (relative prominence of the two syllables of *bamboo* reversed), but not in *oblíque* (first syllable unstressed), as in *oblìque ángle*. The QR constraint, like any OT constraint, is a declarative statement of well-formedness: it cannot be self-limiting so as to only alter the relative prominence of pre-existing foot heads, but simply requires that secondary stress is four syllables from the primary stress, regardless of whether this requires moving stress onto a previously unstressed syllable or not. Therefore, to ensure that, at the Phrase level, the QR constraint does not move stress onto syllables which are not the heads of feet in the individual words involved, or conversely delete the stresses of individual words, QR must interact with constraints like DEP-FootHead ('the input correspondent of an output foot must be a foot head') and MAX-FootHead. This is shown in (9)

(9) Phrasal rhythm with QR[10]

Inputs: *o(blìque) (án)gle*, *(bàm)(bóo) (cúr)tain*	MAX-FootHead	DEP-FootHead	QR
☞a. o(blìque) (án)gle			***
b. (òblique) (án)gle	*!	*	**
a. (băm)(bòo) (cúr)tain			***!
☞ b. (bàm)(bŏo)(cúr)tain			**

Because the QR constraint is not self-limiting so as to only adjust the prominence of pre-existing foot heads, but simply requires secondary stress to be four syllables from the main stress, it can therefore also express the preference for *Èpaminóndas* over *Epàminóndas* in terms of markedness. We can substitute ALIGN(ω, L; Σ, L) (§8.1, §9.1) with QR and achieve the same correct result with respect to underlying foot heads in #LLL words, as shown in (10):

(10) QR and #LLL words

Inputs: */E(pámi)nondas/*, */o(rígi)nality/*, */abracadabra/*	MAX-FootHead	QR
☞a. E(pàmi)(nón)das		**
b. (Èpa)mi(nón)das	*!	*
☞a. o(rìgi)(náli)ty		**
b. (òri)gi(náli)ty	*!	*
a. a(bràca)(dábra)		**!
☞b. (àbra)ca(dábra)		*

Therefore, as with MAX-FootHead>> ALIGN(ω, L; Σ, L), MAX-FootHead must always outrank QR in the stem-level grammar: any foot-head preservation failure is due to the absence of an underlying foot head, rather than anything to do with the constraint ranking (§9.1).

All of the markedness constraints that are crucial to the analysis of left-edge stress preservation on the Stem level have now been established. The constraint

[10] A constraint to ensure the subordination of stresses in the pre-modifiers is excluded here for simplicity.

STRESSIDENT(I) has been proposed to ensure relative prominence preservation. The constraint QR has been proposed which conflicts with both second syllable foot-head preservation and with relative prominence preservation by requiring that secondary stress is ideally four syllables from the word's main stress. Finally, MAX-FtHd has been established as the constraint which ensures solely the preservation of foot-heads.

9.3 Failure of relative prominence preservation

9.3.1 Failure of relative prominence preservation (i): frequency effects

The failure of relative prominence preservation as a result of frequency effects is just like the failure of foot-head preservation as the result of frequency effects presented in §9.1 of this chapter: it fails not because of any variation in the phonology, but because the stress to be preserved is not present in the input.

Assume, for now, a constraint ranking STRESSIDENT(I)>>QR. Under STRESSIDENT(I)>>QR, relative prominence preservation will only fail when the speaker does not recognise that the embedding word contains the embedded word, just as with simple foot-head preservation and MAX-FootHead>> ALIGN(ω, L; Σ, L) in §9.1. As we have argued in this chapter, one possible cause (and the only one addressed in this thesis) of the failure by a speaker to recognise the presence of the embedded word are the frequency effects discussed in chapter 6. This type of preservation failure is shown in (11):

(11) Failure of relative prominence preservation (i)

(a) Successful relative prominence preservation

Input: /*(àm)(bássa)dor-ial*/	STRESSIDENT(I)	QR
a. (àm)(bǎssa)(dóri)al	**!	*
☞ b. (ǎm)(bàssa)(dóri)al	*	**
c. (àmba)ssa(dóri)al	**!*	*

(b) Unsuccessful relative prominence preservation

Input: /*ambassadorial*/	STRESSIDENT(I)	QR
☞ a. (àm)(bǎssa)(dóri)al		*
b. (ǎm)(bàssa)(dóri)al		**!
☞ c. (àmba)ssa(dóri)al		*
Input: /*Luxipalilla*/		
☞a. (Lùx)(ǐpa)(lílla)		*
b. (Lǔx)(ìpa)(lílla)		**!
☞c. (Lùxi)pa(lílla)		*

In (11), relative prominence preservation only occurs when the metrical structure of the embedded word is present in the input – (11a). When the input is unfooted, as for a standard monomorphemic word, (11b), secondary stress falls on the word's initial syllable, not the word's second syllable – candidates a. and c. of both evaluations in (11b). (For both inputs in (11b), there are two equally harmonic candidates: one where the second syllable is completely unstressed (candidates c.), and one where it is the head of a foot and bears tertiary prominence (candidates a.). Which footing is eventually preferred for non-preserving #HLL words depends upon the preferred stress pattern for monomorphemic words. This issue is left for now as it is of some controversy, as shown in §9.4 below.)

9.3.2 Failure of relative prominence preservation (ii): variation in the grammar

In light of the claim that lower perceptual cost is associated with the failure of just relative prominence preservation than of foot-head preservation (§9.2), it is proposed here that speakers may sometimes allow the fulfilment of markedness requirements for eurhythmic stress to take precedence over the desire to preserve relative prominence, even when the relative prominence contour of the embedded word is present in the input. As already argued, the same is not proposed for foot-head preservation: given the suggested greater perceptual cost associated with the failure of foot-head preservation, it is proposed that this phenomenon is inviolable in the grammar – foot-head preservation can only fail as a result of failure to associate embedding and embedded words, such as occurs as a result of frequency effects, and hence variation in the input.

One way of capturing the proposed intrinsically fallible nature of relative prominence preservation in the grammar is to allow the relative prominence preservation constraint, StressIdent(i), and the conflicting markedness constraint, QR, to vary their ranking with respect to one another. One theory which proposes variable constraint ranking is the theory of Partially Ordered Grammar proposed by Anttila (1997, 2002b, 2007) and Anttila & Cho (1998). In a standard OT grammar, constraints are ranked with respect to one another wherever the ranking makes a difference to which the candidate the grammar predicts is the winning output (i.e. where the ranking is crucial), as in A>>B>>C. In a partially ordered grammar, some constraints may be crucially unranked with respect to one another. The grammar consists of 'partial orders' – ranked pairs – of constraints, e.g. {A>>B, A>>C}, rather than a single total order of crucially ranked constraints. These partial orders can then translate into two totally ordered grammars which are subsets of the partially ordered grammar {A>>B, A>>C}: A>>B>>C and A>>C>>B. Phonological variation can be predicted by the grammar because the partially ordered grammar allows two different total orderings of constraints which predict two different winning candidates.

Partially Ordered Grammar is shown to be compatible with Stratal OT by Bermúdez-Otero & McMahon (2006). Bermúdez-Otero & McMahon use a sub-theory of partial ordering, in the form of stratum-internal cophonologies (Anttila, 2002a): different morphological categories subscribe to different total orderings of the stratum's partially ordered grammar.[11] Bermúdez-Otero & McMahon argue that morphologically conditioned partial ordering is required on the Stem level to account for the variable behaviour of extrametricality associated with different adjectival suffixes: syllable extrametricality applies in *ómi<nous>*$_A$, *orígin<al>*$_A$ and *tóle<rant>*$_A$, but only consonant extrametricality applies in *atómi<c>*$_A$, *intrepi<d>*$_A$. Bermúdez-Otero & McMahon propose that this variable behaviour can be captured if the syllable extrametricality constraint, NonFin, is crucially unranked with respect to the constraint which requires syllables to be parsed, Parse-σ, in the stem-level grammar. As shown in (12), this allows Parse-σ and NonFin to be totally ranked in

[11] Strata themselves are also cophonologies, defined by categories like Stem and Word. The difference is that stratal cophonologies are serially ordered.

two ways: NonFin>>Parse-σ, predicting extrametricality in *(ómi)<nous>*, *o(rígi)<nal>* and *(tóle)<rant>*, and the reverse ranking Parse-σ>>NonFin, which predicts non-extrametrical final syllables in *a(tómi)<c>* and *in(trépi)<d>*. (Suffixes are specified as to which of these two cophonologies they subscribe.)

(12) Partial ordering and cophonologies in Bermúdez-Otero & McMahon (2006)

(a) Stem-level partially ordered grammar:

Ft-Bin>>NonFin

{NonFin, Parse-σ}

(b) Cophonology A:

FtBin>>NonFin>>Parse-σ *o(rígi)<nal>*

(c) Cophonology B:

FtBin>>NonFin

Parse-σ>>NonFin *a(tómi)<c>*

Bermúdez-Otero & McMahon's use of morphologically conditioned partial ordering appears to be the most desirable way of dealing with the variable phonological behaviour they observe. The general cophonology approach is preferred to an approach like lexically indexed constraints (e.g. Pater, 2000) as it respects Stratal OT's strict argument that phonology does not refer directly to morphological syntactic information directly (the 'Indirect Reference Hypothesis' (Bermúdez-Otero, in preparation)). Anttila (2002a: 30-33) also shows that cophonologies are preferable to lexically or morphologically indexed constraints on empirical grounds. The argument for a specifically partial ordering theory of cophonologies, cp. Orgun (1996a) and Inkelas (1998), comes from Anttila's (2002a) illustration that partial ordering is restrictive: cophonologies share in common certain phonological generalisations of the grammar, and so can diverge only so much. In sum, the theory of Partially Ordered Grammar is the best way of capturing morphologically conditioned phonological variation in Stratal OT.

Here, it is proposed that Partially Ordered Grammar be extended to the analysis of variable relative prominence preservation. There is no indication that variable relative prominence relative preservation is sensitive to morphology at the Stem level, ruling out a cophonology analysis. Nevertheless, Partially Ordered Grammar has been established in Stratal OT, and so it is logical to propose partial ordering to account for the purely phonological variation of variable relative prominence preservation.[12]

In the partial ordering analysis of the behaviour of foot-head and relative prominence preservation, three constraints will be assumed: MAX-FootHead, STRESSIDENT(I) and QR. To keep the analysis as simple as possible, other constraints mentioned so far in this chapter that are immediately superfluous, e.g. ALIGN(ω, L; Σ, L), are excluded.

In §9.2.2, we established that the invariant stem-level ranking of MAX-FootHead>>QR ensures foot-head preservation whenever a foot-head is present in the input. This ranking is therefore fixed in all stem-level total rankings, and is thus a phonological regularity that holds over all stems. Conversely, as noted earlier in the present section, the relative prominence preservation constraint, STRESSIDENT(I), and QR must be variably ranked.

With three constraints, the factorial typology predicts six (3!) possible stem-level total rankings of the three constraints MAX-FootHead, STRESSIDENT(I) and QR. However, we have argued that there is a total ranking that must hold in all the sub-grammars – MAX-FootHead >> QR – which leaves three possible stem-level total rankings:

(13) Stem-level partial and total rankings for relative prominence and foot-head preservation

(a) Stem-level partially ordered grammar:

MAX-FootHead>>QR

{QR, STRESSIDENT(I)}

[12] Additionally, Anttila (2002b, 2007) argues that partial ordering is preferable to the multiple grammars approach of Kroch (1989) which could be employed here

(b) Inventory of totally-ranked sub-grammars

Total ranking 1: MAX-FootHead>>QR>>STRESSIDENT(I)

Total ranking 2: STRESSIDENT(I)>>MAX-FootHead>>QR

Total ranking 3: MAX-FootHead>>STRESSIDENT(I)>>QR

In practice, STRESSIDENT(I) and MAX-FootHead are not constraints which are in conflict, and so total rankings 2 and 3 in (13b) will predict the same outcome.

The results of all three total rankings when the metrical structure of the embedded word is present in the underlying representation are shown in (14):

(14) Behaviour of foot-head versus relative prominence preservation

(a) Total ranking 1: MAX-FootHead>> QR>>STRESSIDENT(I)

Inputs: /*o(rígi)nal-ity*/, /*(àn)(tíci)pate-ion*/	MAX-FootHead	QR	STRESS IDENT(I)
☞a. o(rìgi)|nálity		**	*
b. (òri)gi|nálity	*!	*	***
a. (ăn)(tìci)|pátion		**!	*
☞b. (àn)(tĭci)|pátion		*	**
c. (ànti)ci|pátion	*!	*	***

(b) Total ranking 2: STRESSIDENT(I)>>MAX-FootHead>>QR

Inputs: /*o(rígi)nal-ity*/, /*(àn)(tíci)pate-ion*/	STRESS IDENT(I)	MAX-FootHead	QR
☞a. o(rìgi)|nálity	*		**
b. (òri)gi|nálity	*!**	*	*
☞a. (ăn)(tìci)|pátion	*		**
b. (àn)(tĭci)|pátion	**!		*
c. (ànti)ci|pátion	**!*	*	*

(c) Total ranking 3: MAX-FootHead>>STRESSIDENT(I)>>QR

Inputs: */o(rígi)nal-ity/*, */(àn)(tíci)pate-ion/*	MAX-FootHead	STRESS IDENT(I)	QR
☞a. o(rìgi)│nálity		*	**
b. (òri)gi│nálity	*!	***	*
☞ a. (ăn)(tìci)│pátion		*	**
b. (àn)(tĭci)│pátion		**!	*
c. (ànti)ci│pátion	*!	**!*	*

In (14), all three total rankings predict successful foot-head preservation, as we want. However, relative prominence preservation in *anticipation* does not enjoy the same 100%-success in (14), even though *(àn)(tíci)pate* is always visible in *anticipation*'s underlying representation: relative prominence preservation is successful in (14b, c), but fails in (14a). Partially Ordered Grammar can therefore model the type of result we want: in the absence of variation in the input, foot-head preservation is inviolable, but relative prominence preservation sometimes fails.

Anttila (2002b, 2007) shows that it is possible to make quantitative predictions from partially ordered grammars: the proportion of total rankings which predicts a particular output should correspond to the probability of that output's real-world occurrence. For the present analysis, this means that, as successful relative prominence preservation is predicted by 2/3 of our total rankings – (14b, c) – and failure by 1/3 – (14a), relative prominence preservation should fail 33.3% of the time, independent of any variation in the input that could be caused by frequency effects or other variables which may cause variation in the input.

Unfortunately, applying this type of quantitative analysis to the left-edge stress preservation data presented in this thesis is problematic. We know that the proportion of tokens which exhibit relative prominence preservation failure are also affected by another factor, albeit comparatively minor in its effect, aside from being candidates for relative prominence preservation: word-frequency effects (§6.5.1). Even worse, the logistic regression analysis presented in §6.5.1 indicates that there are additionally unidentified factors which are affecting the success of second-syllable left-edge stress preservation (R^2 is nowhere near 100%).[13] To my knowledge, identifying the proportion of words which have relative prominence preservation failure as a result of the assumed interaction between STRESSIDENT(I) and QR alone will require further data collection in controlled conditions, with factors like frequency kept constant from the outset. Nevertheless, it is worth pointing out that the analysis in terms of partial ordering predicts that relative prominence preservation will be successful most of the time. This does not conflict with the observations presented in chapter 4, where it was also observed that relative prominence preservation was successful most of the time. Inasmuch as the partial ordering analysis does not predict failure of relative prominence preservation the majority of the time, it is in keeping with the empirical observations of this thesis.

9.3.2.1 Prefixes and cophonologies

There is another phonological phenomenon discussed in this thesis which seems to be amenable to a partial ordering analysis: stem-level Latinate prefixations. This phenomenon specifically merits an analysis in terms of the stratum-internal cophonologies proposed by Bermúdez-Otero & McMahon (2006).

It has been argued in this thesis that the footing of #HLL and #HHL second-syllable preserving words is #(H)(LL)/ #(H)(HL), with the word-initial syllables stressed. However, as discussed in §6.5.4.2, these heavy, word-initial syllables nevertheless seem to be sometimes unstressed when they correspond to an opaque Latinate prefix: e.g. ***con****glomeration*, which has both [kɒn] and [kən], cp. non-prefixed *participation* which has consistently [pɑː].

Both *c*[ə]*n(glome)ration* and (*p*[ɑː])(*tici)pation* could be handled on the Stem level if we propose that two constraints, PARSE-σ and *CLASH (the latter meaning 'no stresses on adjacent syllables'), are crucially unranked in the stem-level phonology. We can then have two stem-level cophonologies: one for words containing obscured Latinate prefixes, where *CLASH>>PARSE-σ rules out **(con)(glome)ration* in favour of *con(glome)ration* and prepares the way for initial-syllable vowel reduction, and

[13] The figure we got in our linear regression analysis for initial-syllable weight in §6.5.1 is an effect size: it tells us how much of the stress preservation behaviour initial-syllable weight can account for, which is not the same as how many words vary or not as a result of the effect of initial-syllable weight.

one for all other words with the opposite ranking, PARSE-σ>>*CLASH, which prefers *(par)(tici)pation* to **par(tici)pation*. This is shown in (15):

(15) Cophonologies for bound Latinate prefixes

{*CLASH, PARSE-σ}

Cophonology A: stems containing bound Latinate prefixes

*CLASH>>PARSE-σ

Input: /*con(glóme)rate-ion*/	*CLASH	PARSE-σ
a. (còn)(glòme) | ration	*!	
☞b. con(glòme) | ration		*

Cophonology B: other stems

PARSE-σ>>*CLASH

Input: /(pàr)(tíci)pate-ion/	PARSE-σ	*CLASH
☞a. (pàr)(tìci) | pation		*
b. par(tìci) | pation	*!	

9.4 #HHL and #HLL monomorphemic and bound-root base words

The analysis of relative prominence preservation and frequency in §9.3.1 was indeterminate as to the stress pattern for #HLL words where there is no metrification present in the input, as shown in (16) (repeated from (11)):

(16) #HLL words without relative prominence preservation

Input: /*ambassadorial*/	QR
☞ a. (àm)(bǎssa)(dóri)al	*
b. (ǎm)(bàssa)(dóri)al	**!
☞ c. (àmba)ssa(dóri)al	*
Input: /*Luxipalilla*/	
☞a. (Lùx)(ǐpa)(lílla)	*
b. (Lǔx)(ìpa)(lílla)	**!
☞c. (Lùxi)pa(lílla)	*

As can be seen in (16), QR is concerned only with the placement of secondary stress, not with the location of lower level stress: #(H̀)(LL) and #(H̀L)L are equally acceptable non-prominence-preserving metrifications. One would expect that this indeterminacy could simply be solved by adding more constraints to the ranking. However, in this section, it is argued that, although #(HL)L appears to be the more appropriate metrification based upon observations of the data from chapter 4, formalising this in constraints is problematic.

In §3.1, the general opinion in the literature was shown to be that the default footing of #HLL words is #(HL)L, accounting for invariant secondary stress in most monomorphemic #HLL words, e.g. *(Lùxi)pa(lílla)*. Second syllables may only exceptionally be the heads of feet, i.e. #(H)(LL), so as to account for variable secondary stress placement in examples like *Dòdecanésian~Dodècanésian*. The data and analysis in chapter 4 appeared to confirm the argument that the default footing is #(HL)L: the great majority of #HLL monomorphemic and bound-root base words had invariant, initial-syllable secondary stress, with a considerable proportion of these words exhibiting vowel reduction in their second syllable. Both observations would follow if these words were metrified #(HL)L. The footing for #HHL monomorphemic or bound-root base words is altogether more difficult to discern. The consensus in past literature is that the footing is #(H)(HL) (Kiparsky, 1979; Halle & Vergnaud, 1987a; Hammond, 1989; Pater, 1995, 2000). The data presented in chapter 4 was unable to argue for or against this claim, as so few tokens of #HHL monomorphemic or bound-root base words were found. In this situation, it seems only sensible to follow past literature and assume the footing #(H)(HL).

A constraint ranking therefore needs to be found that predicts the metrifications #(HL)L and #(H)(HL) in the absence of preserving or exceptional underlying second syllable stress. The most exhaustive analysis of English secondary stress in OT to date is Pater (1995, 2000). It is therefore very interesting that this analysis is unable to predict both #(HL)L and #(H)(HL).

Pater (1995: 39, 2000: 269-70) shows that his constraint hierarchy predicts the incorrect footing for #HLL words: #(H)(LL) *(Lùx)(ìpa)(lílla)*, not #(HL)L

(Lùxi)pa(lílla) ('The *Luxipalilla* problem').[14] In (17), a revised version of Pater's constraint ranking is given. This ranking has been revised only to exclude pre-tonic stress in all words so as to rule out pre-tonic stress preservation (the rankings of WSP and *CLASH-HEAD have been reversed, shown in (19) of chapter 2); the ranking makes the same predictions with respect to #HLL:

(17) Pater's (2000) ranking cannot predict #(HL)L

ALIGN(Σ, L, ω, L) ('ALIGN-L') = align all feet with the left edge of the prosodic word

Input: /*Luxipalilla*/	FTBIN	PARSE-σ	*CLASH-HEAD	STRESS IDENT	WSP	ALIGN-LEFT
💣 a. (Lùx)(ìpa)∣lilla						*
☹ b. (Lùxipa)∣lilla	*!					
☹ c. (Lùxi)pa∣lilla		*!				
Input: /*Ticonderoga*/						
☞ a. (Tì)(cònde)∣roga						*
b. (Tìconde)∣roga	*!				*	
c. (Tìcon)de∣roga		*!			*	

This analysis predicts the right result for #HHL words like *Ticonderoga*, but the relatively high ranking of FTBIN and PARSE-σ with respect to ALIGN-L (=ALLFEETLEFT: 'align all feet with the left edge of the prosodic word'; Pater (2000: 243)) forces the incorrect default footing *#(H)(LL) for #HLL words: both #(HL)L, and a potential acceptable alternative, #(HLL),[15] are ruled out.

Clearly, Pater's constraint ranking needs some revision if it is to predict #(HL)L rather than #(H)(LL). However, this is not simple, as the existing problematic rankings are well-motivated in other English words. First, the ranking of PARSE-σ above ALIGN-L, which rules out *(Luxi)pa /lilla* in favour of *(Lux)(ipa) /lilla*, is required to avoid long strings of unstressed syllables:

> McCarthy & Prince (1993b) show that iterative stress placement requires the dominance of PARSE-σ, and as examples like *Àpalàchicóla* show, English

[14] In the pronouncing dictionaries used in this thesis, the vowels of the third syllables of #HLL words are also shown as prone to reduction, ruling out Pater's (1995: 270) suggestion that the pre-tonic disyllabic feet his analysis predicts could be iambic, i.e. *(Lùx)(i.pă)lílla*.

[15] FTBIN functions as an upper as well as lower limit in Pater (2000): trisyllabic feet are excluded.

> does have iterative secondary stress […]. ALIGN-L is fully satisfied only if there is a single foot at the left edge, but PARSE-σ, which demands all syllables belong to feet, requires medial footing (Pater, 2000: 244).

We therefore apparently have to retain PARSE-σ>> ALIGN-L to correctly predict the stress pattern of *Àpalàchicóla*, even though this ranking excludes desirable candidate *(Luxi)pa /lilla* in (17).

(Luxipa) /lilla is also ruled out in favour of *(Lux)(ipa) /lilla*. This is because, by transitivity of ranking, FTBIN must be ranked above ALIGN-LEFT in English. FTBIN must be ranked above STRESSIDENT to ensure that preservation does not occur on the initial syllables of words like *atómic* (*átom* → **àtómic*). In turn, Pater's STRESSIDENT must be ranked above ALIGN-LEFT to ensure that second-syllable foot-head preservation is successful, as in *oríginal* → *orìginálity*. Therefore, by transitivity, FTBIN>>ALIGN-LEFT. If, as argued by Pater (1995: 39), FTBIN excludes trisyllabic feet, then FTBIN>>ALIGN-LEFT excludes the other footing which leaves the second syllable unstressed, as desired: b. *(Luxipa) /lilla*, in (17).

It seems, then, that Pater's (1995, 2000) analysis must be seriously revised. As will be seen by the rest of this section, this is not successfully done here. Nevertheless, the following exploration gives some food for thought which will assist future analyses of pre-tonic #HLL sequences.

One area which most certainly merits reconsideration is Pater's handling of quantity sensitive stress. An obvious factor which corresponds to the different footings of #HLL and #HHL monomorphemic and bound-root base words is the weight of the second syllable: this syllable is stressed when it is heavy – #(H)(HL) – and unstressed when it is light – #(HL)L. (It will be assumed, for now, that we do not want to introduce ternary feet into our analysis, and will not consider #(HLL) as in *(Lùxipa)(lílla)*.) As it stands, Pater's analysis cannot take this apparent weight distinction into account: the relatively high ranking of PARSE-σ ensures that #HLL and #HHL sequences both have the same footing, #(σ)(σσ). One way to amend this would be to rank a constraint which prefers light syllables to be unstressed above PARSE-σ. Therefore, in addition to Pater's (2000) WEIGHT-TO-STRESS (WSP), which requires that all heavy syllables are stressed, we must include the opposite constraint, STRESS-TO-WEIGHT, which requires that **only** heavy syllables are stressed (e.g.

Kager, 1999).[16] By ranking STRESS-TO-WEIGHT above PARSE-σ, the analysis predicts the right results for #HLL and #HHL words:

(18) Introduction of STRESS-TO-WEIGHT

Input: /*Luxipalilla*/	STRESS-TO-WEIGHT	PARSE-σ
a. (Lux)(ipa) | lilla	*!	
☞b. (Luxi)pa | lilla		*
Input: /*Ticonderoga*/		
☞a. (Ti)(conde) | roga		
b. (Ticon)de | roga		*!

However, the introduction of STRESS-TO-WEIGHT has other less positive implications: this analysis now predicts the wrong outcome for long words like *Àpalàchicóla*, as shown in (19):

(19) Incorrect prediction of STRESS-TO-WEIGHT>>PARSE-σ

Input: /*Apalachicola*/	STRESS-TO-WEIGHT	PARSE-σ
☹ a. (Àpa)(làchi) | cola	*!	
💣 b. Apalachi | cola		****

We therefore need a constraint which can prevent incorrect outcomes like that in (19), while retaining the ranking STRESS-TO-WEIGHT>>PARSE-σ necessary to predict #(HL)L, not *#(H)(LL).

The solution to the **Apalachi* /*cola* problem appears to be the constraint *LAPSE:

(20) *LAPSE (Elenbaas & Kager, 1999)

Every weak beat must be adjacent to a strong beat or the word edge.

*LAPSE prohibits strings of more than two consecutive unstressed syllables in words, forcing *(Àpa)(làchi)cóla*, but with no violations incurred by *(Lùxi)palilla*.[17] *LAPSE

[16] Ricardo Bermúdez-Otero (personal communication) suggests that a putative #(H)(HL) vs. #(HLL metrification asymmetry could be attributable to STRESS-TO-WEIGHT.

[17] In the (largely American) generative literature on Metrical Phonology with which I am familiar, there is a strong consensus that no sequences of more than two unstressed syllables are permitted.

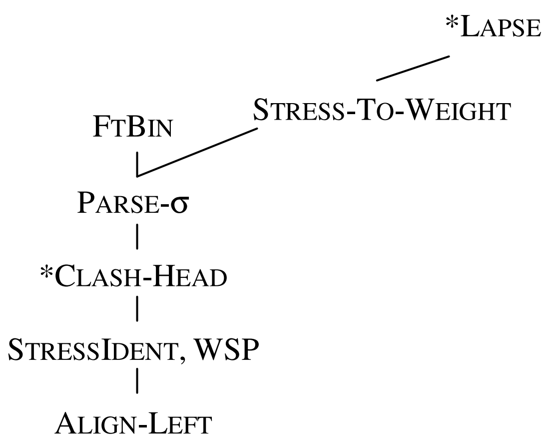
must be ranked above STRESS-TO-WEIGHT, as shown in (21):

(21)

Input: /*Luxipalilla*/	*LAPSE	STRESS-TO-WEIGHT	PARSE-σ
a. (Lux)(ipa)⎸lilla		*!	
☞b. (Luxi)pa⎸lilla			*
Input: /*Ticonderoga*/			
☞a. (Ti)(conde)⎸roga			
b. (Ticon)de⎸roga			*!
Input: /*Apalachicola*/			
☞a. (Àpa)(làchi)⎸cola		**	
b. (Àpa)lachi⎸cola	*!	*	**
c. Apalachi⎸cola	*!		****

The Hasse diagram for the new constraint hierarchy is given in (22):

(22)



Unfortunately, this ranking by no means solves all of the problems, as shown in (23):

However, French corpus-based work by Lionel Guierre argues for #2000- strings (Nicolas Ballier, personal communication).

(23) Problems of the new hierarchy

Input: /*Alexander*/	STRESS-TO-WGT	PARSE-σ	STRESS IDENT	WSP	ALIGN-LEFT
☹ a. (Àlex) | ander	*!			*	
💣 b. A(lèx) | ander		*			*
Input: /*(àm)(bássador)-ial*/					
☹ a. (ăm)(bàssa) | dorial	*!		*		*
💣 b. (àmba)ssa | dorial		*	**		

By having STRESS-TO-WEIGHT ranked above PARSE-σ and ALIGN-LEFT, the ranking predicts the incorrect parsing for *Alexander*: **A(lèx)(án)der*, not (*Àlex)(ánder)*. It must also be considered where the preservation constraints established earlier in this chapter fit into the hierarchy. For example, Pater's STRESSIDENT is currently ranked below STRESS-TO-WEIGHT, with the damning consequence that stress will never be preserved in #HLL words like *ambassadorial*.

Finally, one other potential route by which the *Luxipalilla* problem could be resolved, without devising a whole new hierarchy from scratch, is to reconsider the definition of FTBIN. As it stands, FTBIN is also a maximal limit which rules out ternary feet, which, high-ranking as it is in Pater's analysis, rules out the other potentially correct footing of #HLL – #(HLL). If FTBIN did not function as an upper limit in this way, then #(HLL) *(Lùxipa) /lilla* would be permitted, which has the desired unstressed second syllable. Unfortunately, this causes major problems for second syllable foot-head preservation, shown in (24):

(24) Revised FTBIN

FTBIN(I) – no upper binary limit upon foot size

Input: /*i(má.gine)-ation*/	FTBIN(I)	PARSE -σ	STRESS IDENT
☹ a. i(mà.gi)nation		*!	
💣 b. (ì.ma.gi)nation			*

In short, as the fairly exhaustive adjustments to Pater's analysis made here show, there is unlikely to be a way of predicting the desired monomorphemic and bound-root base metrification #(HL)L/#(HLL) while still retaining most of Pater's (1995, 2000) constraint hierarchy. A new working from scratch may be required, which, like Pater's original analysis, will additionally need to take into account the secondary stress patterns of diverse English words; this is clearly a project too extensive for this thesis.[18]

9.5 Conclusion

In this chapter, both the constraints and the method of constraint ranking needed to capture foot-head and relative prominence preservation and their different degrees of success have been presented. The analysis has required the introduction of new constraints, and also Anttila's theory of Partially Ordered Grammar. Problems with Pater's (1995, 2000) constraint hierarchy for English secondary stress were also exemplified, to which no solutions were found.

[18] Hammond (1999a: 316) does propose an analysis for #HLL words, but it seems less than ideal. The only way the metrification #(HL)L can be assured is by faithfulness to an underlying lexical accent (≈ foot head in this chapter); otherwise, the pattern #H(LL) is predicted. Given that #(HL)L is argued to be the default stress pattern here, Hammond's solution seems to have it the wrong way around.

Summary and conclusions

10.0 Introduction

The summary and conclusions of the thesis are presented in this chapter. A brief chapter summary is presented in §10.1, before the findings of the thesis and their implications are discussed in more detail in §10.2. Speculations concerning future research are then made in §10.3.

10.1 Chapter summary

Following the general introduction of the thesis, its theoretical context was presented in chapter 1. This chapter reviewed the formal proposals that are central to the analysis of English stress preservation presented here: the cycle, strata, and Optimality Theory. This chapter showed that Bermúdez-Otero's (in preparation) models of Stratal OT draws upon all of these proposals, but also fundamentally differs from them in certain ways. A particular new proposal outlined here that becomes central in the analysis of weak stress preservation later in the thesis is 'fake cyclicity'.

Chapters 2 and 3 presented the arguments which question the status of the two main types of weak stress preservation in English. In chapter 2, the very existence of the first type of weak stress preservation – pre-tonic stress preservation – was rejected. The outcome of chapter 3 was not so extreme: it did not deny the existence of left-edge stress preservation altogether. This chapter did, however, show that there is controversy over the precise status of left-edge stress preservation which could only be resolved by new empirical investigation.

The required empirical investigation into left-edge stress preservation was undertaken in chapters 4 and 5, for which data was taken from English pronouncing dictionaries. In chapter 4, the data for relative prominence preservation, the first sub-type of left-edge stress preservation, was examined. There appeared to be evidence for relative prominence preservation, but also an indication that the process is not consistently successful. Data relevant to just the second sub-type of left-edge stress preservation, left-edge foot-head preservation, was examined in chapter 5. Again, although there were instances of preservation failure, the data argued strongly in favour of left-edge stress preservation, and statistical evidence was given to this effect.

Chapter 6 was a data analysis chapter that attempted to establish one reason why left-edge stress preservation observes inconsistent success. Previous work indicated that word frequency might reasonably affect the success of stress preservation, and so this variable was examined. Save a single statistical analysis in §6.5.1 which dealt with words that were candidates for just left-edge foot-head preservation, the data only permitted an analysis of relative prominence preservation with respect to word frequency, but the result was encouraging: word frequency does significantly predict the success of relative prominence preservation. (The analysis in §6.5.1 indicated the same effect for words which were candidates for just foot-head preservation.)

In chapters 7, 8 and 9, the theoretical implications of the empirical investigation and analysis were considered. Chapter 7 presented the general problems the existence of a phenomenon like English stress preservation poses for OT, and the two possible solutions to it: Output-Output Correspondence Theory, and Stratal OT. Significant problems of Output-Output Correspondence Theory were identified; Stratal OT was shown to be altogether more promising. In chapter 8, it was shown how Stratal OT handles the probabilistic nature of left-edge stress preservation: 'fake cyclicity'. Finally, chapter 9 presented the OT constraints and rankings required to handle left-edge foot-head and relative prominence preservation respectively on the Stem level.

10.2 Thesis summary, conclusions and implications

This empirical investigations carried out in this thesis have revealed two important new observations:

(i) There is left-edge stress preservation in English
(ii) English left-edge stress preservation is probabilistic.

Up until now, the existence or absence of left-edge stress preservation has simply been assumed without any serious empirical investigation and analysis being undertaken to support such a conclusion (chapter 3). The present thesis has aimed to rectify this situation, not least because, as will be discussed shortly, the nature of a morphophonological identity effect like stress preservation has the potential to tell us much about the operation of the interface between morphology and phonology.

The empirical investigations have been based upon data taken from two English pronouncing dictionaries, Jones (2003) and Wells (2000). The data consists of comprehensive lists of words from both dictionaries which have three pre-tonic syllables, some of which are candidates for initial- or second-syllable preservation, and others which are not preservation candidates and therefore serve as control data. From analysis of the data, it has been shown that there is statistical evidence for relative prominence preservation and foot-head preservation. The result for relative prominence preservation is particularly interesting, as this phenomenon has been given little attention since it was proposed by Kiparsky (1979) and then rejected by Halle & Vergnaud (1987a).

The precise nature of left-edge stress preservation has also been considered. Instances of failure of both relative prominence and foot-head preservation have been noted in past literature (chapter 3), and were similarly present in the data collected from pronouncing dictionaries (chapters 4 and 5). In chapter 6, it was hypothesised that this failure could be attributable, at least in part, to word-frequency effects. A particular hypothesis which was explored was Hay's (2001, 2003) proposal for word-specific relative frequency effects. The data for relative prominence preservation was analysed, and the results supported the relative frequency hypothesis: embedding words which are more frequent than their embedded words are less likely to preserve stress than embedding words which are less frequent than their embedded words. There were also effects of absolute frequency upon relative prominence preservation, which Hay (2003) argues to be symptomatic of relative frequency effects.

The frequency analyses presented in the thesis focused upon relative prominence preservation; due to insufficient data in certain categories, little attention was given to left-edge foot-head preservation. However, it is both extremely plausible and logical that the same type of frequency effects will affect for instances of left-edge foot-head preservation too, and this has been assumed in the analysis in the remainder of the thesis. The single frequency analysis which does consider words that are candidates for just left-edge foot-head preservation, given in §6.5.1, supports this hypothesis: the same types of absolute frequency effects are observed as are documented for relative prominence preservation in the rest of the chapter.

The existence and nature of left-edge stress preservation has been focused upon in the empirical investigations here, as the status of the other type of weak stress preservation, pre-tonic stress preservation, is seriously in doubt. In chapter 2, it was shown that the assumption of a symmetrical relationship between vowel reduction and stress, which provides the only empirical argument for pre-tonic stress preservation, incurs considerable problems for the English stress system: stressing of extrametrical syllables; introduction of word-initial degenerate feet; the stipulation that certain stress rules are sensitive to coda-place (coronal versus non-coronal); and, finally, the rejection of the final-consonant extrametricality generalisation for English. It was argued that these problems could be avoided if an asymmetrical relationship between vowel reduction and stress was adopted, but that this argument removes the case for pre-tonic stress preservation.

Nevertheless, it has been argued that vocabulary which has been traditionally argued to display pre-tonic stress preservation does still display preservation – it is just of vowel quality, not stress (§2.5, §6.1.1). Vowel quality preservation, like stress preservation, is a morphophonological identity effect. Once more, there are early indications that vowel quality preservation is probabilistically dependent upon word frequency, like left-edge stress preservation (§6.1.1).

The formal implications of the preservation phenomena considered in this thesis are therefore twofold. First, phonological theory must be able to handle morphophonological identity effects. This has been shown to have major implications for OT, requiring a rejection of the generally assumed monostratal model of OT in favour of a non-standard, multi-stratal model – Stratal OT (chapter 7). The second implication is that a phonological theory must be able to capture the demonstrated probabilistic nature of these morphophonological identity effects. This has been shown to be a problem for the cycle (chapter 8). Since SPE, it has been recognised that the cycle is good at handling weak stress preservation (so much so, that it has even been imported in disguise into monostratal OT – Benua (1997), discussed in chapter 7). Here, it has been shown that the cycle cannot explain the probabilistic nature of left-edge stress preservation. Fake cyclicity, its replacement, suffers no such problem: preservation is enforced through blocking, an independently established psycholinguistic phenomenon which depends probabilistically upon token word frequency.

Fake cyclicity constitutes a radical departure from previous stratal models. Models of LPM (e.g. Kiparsky, 1982; Giegerich, 1999) have proposed that stratum one is internally cyclic in order to handle weak stress preservation. Fake cyclicity rejects this stratum-internal cycle, proposing that every output of the stratum can be produced by a single pass through the stratum's phonology. Preserved stress which appears to be the result of cyclic application is instead the result of a combination of lexical storage, redundancy rules and blocking. The fake cyclicity analysis allows Stratal OT to propose a more restrictive theory of phonological opacity, and resolve phonological domain mismatches on stratum one (§8.1). However, fake cyclicity is also psycholinguistically well-motivated, and vital for the analysis of stress preservation failure on stratum one. Fake cyclicity's resemblance to the dual-route model of lexical access means it can predict preservation failure; the stratum-internal cycle cannot. Fake cyclicity's crucial assumption that all stratum one outputs are stored in the speaker's memory, along with their predictable metrical structure, is also supported by more general psycholinguistic arguments about the balance of storage and computation, and the tip-of-the-tongue phenomenon.

In conclusion, this thesis constitutes the most rigorous investigation of English weak stress preservation to date, and has shown it to support Bermúdez-Otero's (in preparation) particular model of Stratal OT, especially this model's formulation of stratum one. The thesis has also indicated the need for further empirical investigation, outlined shortly in §10.3. Before this, it is considered whether the findings can solve a problem of Giegerich's (1994, 1999) model of LPM that was discussed in chapter 1.

10.2.1 Implications for Giegerich's (1994, 1999) model of LPM

As noted in §1.2.1.2, Giegerich's (1994, 1999) models of LPM has a particular problem with stratum one's internal cycle. In Giegerich's model, lexical category assignment is assigned upon exit from the highest stratum, via the Root-to-Word rule. Although motivated for other sound reasons, the Root-to-Word rule confounds stratum one cyclic stress assignment. English primary word stress is well known to be sensitive to lexical category (Hayes, 1982), meaning that primary stress cannot be assigned until after Root-to-Word, on an item's exit from stratum one. The implication of this is that there can be no stratum one cyclic stress assignment: once a lexical item is a word, it cannot pass through the Root level's phonology. So, for example, the stress of *oríginal*$_A$ could not be preserved in the root-level formation *originality* because the word *oríginal*$_A$ cannot be an input to root-level phonology. In this thesis, it has been proposed that stratum one is internally noncyclic. It is therefore worth considering whether this proposal for stratum one noncyclicity can solve the problem of Giegerich's stratum one.

Under fake cyclicity, the operation of weak stress preservation is as shown in figure 1: [1]

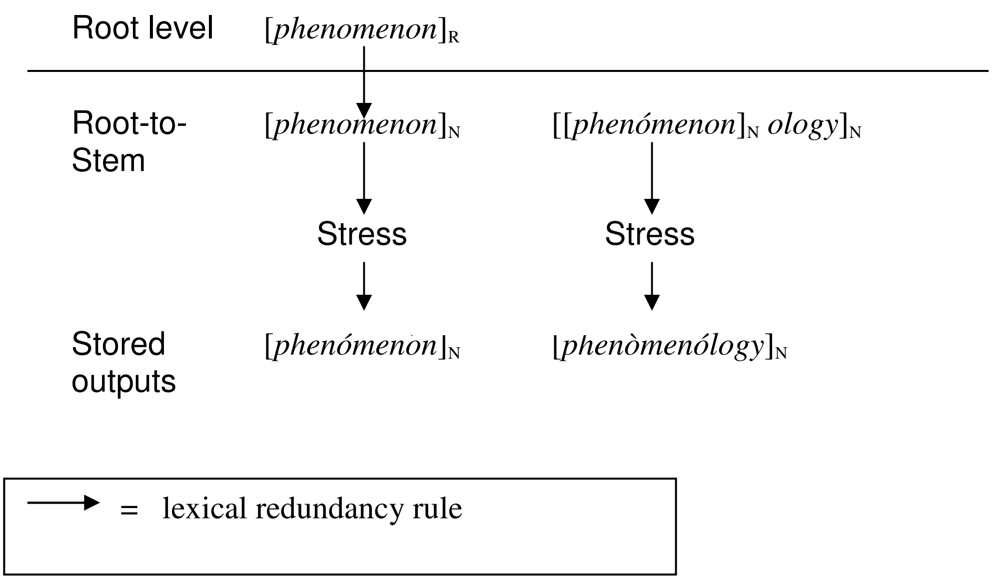


Figure 1: weak stress preservation in Bermúdez-Otero (in preparation)

In Giegerich's model, which rejects the Stem level, the structure is as shown in figure 2:

[1] Bermúdez-Otero (in preparation) argues that Root-to-Stem conversion is Stem level. In contrast, Giegerich (1999) proposes that Root-to-Word is Root level.

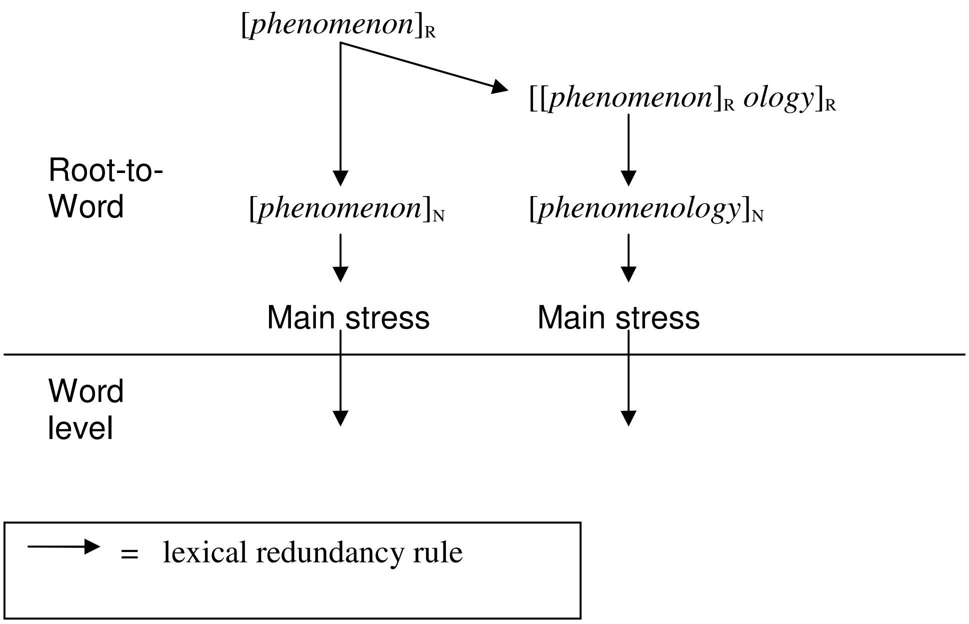


Figure 2: stratum one in Giegerich (1994, 1999)

There is a crucial difference between figures 1 and 2 with respect to stress preservation. In figure 1, *phenómenon* already has primary stress when it is suffixed with -*ology*, allowing preservation to take place. It is possible for *phenómenon* to bear primary stress because it is a stem, and so the lexical category that predicts primary stress placement is known. In contrast, in figure 2, it is the unmetrified root *phenomenon* that is suffixed with -*ology*. The root is necessarily unmarked for primary stress because its lexical category is not known. Clearly, noncyclic stress assignment can only predict the correct results for weak stress preservation if the stress of the embedded form is present in the 'underlying representation' of the embedding form; for this to happen, the embedded form must be marked for lexical category, something prohibited in Giegerich's model. Therefore, with respect to weak stress preservation, the Stem level seems to be unavoidable in English. A similar conclusion is reached by Hurrell (2001: 281) in her consideration of the Root-to-Word problem in Giegerich's model, her final solution suggesting "lexically categorised roots".

As far as weak stress preservation in Giegerich's model is concerned, the only other possible solution would be to show that English primary stress is not sensitive to lexical category. However, even if this is the case, assigning stress to roots is likely to be contentious. Giegerich argues for the existence of the Root stratum on the basis of the existence of bound roots like *matern-*. It is completely stipulative to propose any stress for such forms, given that they can never occur independently. If stress cannot reasonably be assigned to prototypical roots, it seems wrong to assign stress on the Root stratum at all.

10.3 Speculations concerning future research

This thesis has opened up some theoretical issues that will require future consideration. For example, the discussion of pre-tonic preservation in chapter 2 indicated the need for a fully-fleshed out theoretical analysis of pre-tonic vowel quality preservation, as this is not an issue dealt with by Burzio (1994, 2002, in press) in his acoustic-perceptual analysis of pre-tonic vowel reduction. Later in the thesis, in chapter 9, it was shown that it still remains to find the overall ranking of stratum one metrical constraints, as those proposed so far predict the wrong stress patterns for one or more groups of words. However, perhaps the most interesting areas for future research will relate to the nature of stress preservation itself: the empirical investigations carried out in this thesis have provided strong evidence for Bermúdez-Otero's model of Stratal OT, but they have also indicated the need for much further empirical research into English stress preservation and related phenomena.

The argument for fake cyclicity has been made on the basis of Hay's (2001, 2003) relative frequency hypothesis. The relative frequency argument allows preservation failure to be explained in terms of blocking failure: the less frequent the embedded word, the lower its resting-activation level, and so the lower its ability to block the default, non-preserving stress pattern. However, relative frequency is by no means the only type of frequency effect that has been argued for in psycholinguistic literature. In chapters 2 and 6 of this thesis, for example, we saw that Hammond has argued for an opposite effect of the frequency of the embedded word with respect to pre-tonic vowel quality preservation: the more, not less, frequent this word is, the more likely preservation is to fail. Hammond's frequency argument certainly does not fit with the blocking argument of preservation failure upon which fake cyclicity crucially relies. Hammond's results and interpretation were shown to be questionable in chapter 6, removing any immediate threat to the fake cyclicity analysis. However, if a new empirical investigation and analysis shows that Hammond's cumulative frequency effect does indeed exist with respect to pre-tonic vowel quality preservation, then an explanation will be required that will allow both cumulative and relative frequency effects to exist in a single grammar.

One speculative solution may come from arguing that different frequency effects apply in different phonological domains. Pre-tonic vowel quality preservation involves vowel reduction, and it has been argued that vowel reduction is a late phonological process in English (SPE: 113; Halle & Vergnaud, 1987a: 240), applying after all stress assignment. This is interesting because Hammond makes another argument for cumulative frequency effects, from phrasal rhythm: high frequency pre-modifiers (e.g. *nàíve*) undergo phrasal rhythm more easily than low frequency ones (AmEng *òbése*) (hence *obèse chíld* but ***nà**ive fríend*), and high-frequency phrases themselves also undergo rhythm more than low-frequency phrases (***ùn**fit móther* but *unfìt fáther*) (Hammond, 1999b; Hicks et al., 2000). It may therefore be that cumulative frequency effects hold at the phrase level, applying to vowel reduction and phrasal rhythm, but relative frequency effects apply on the Stem level, the home of blocking. Of course, it would then have to be considered why there is such domain circumscription of frequency effects: for example, are the mental representations and/or methods of lexical access for phrasal formations different to those of stems? (It should also be noted that speakers can only have knowledge of the frequencies of phrases like *unfit mother* if these phrases are stored whole somewhere in their memory. This obviously has implications for the present proposal that it is specifically the outputs of the Stem level that are stored in a speaker's permanent lexicon.)

In the context of the frequency analysis, issues also arose with determining the source of preserved stress. The phonological cycle and analogues are very specific about the source of preserved stress – the immediately embedded, strictly local word. From a psycholinguistic perspective, this is too restrictive: instances have already been discussed where the phonology of a more-deeply embedded word may or certainly affects that of the embedding word: e.g. ***tó**tal~totálity~**tò**talitárian*;

*c***[aɪ]***cle~c*[ɪ]*clic~c***[aɪ]***clicity* (see §8.3.2). However, no statistical relationships between embedding words and their more-deeply embedded words have been given in this thesis. This would, therefore, be a desirable topic for future research (although it is possible that, with respect to the stress preservation data presented here, controlling for suffixation may leave too little data for statistical analysis). Such an analysis could also shed further light upon relative frequency effects, as the focus of relative frequency analyses so far has been the relationships between derived words and their morphologically simple bases (e.g. Hay & Baayen, 2002: e.n. 2).

Other issues which need to be considered are the role of semantic and phonological similarity in preservation. In chapters 4 and 5, words were treated as being candidates for preservation even if the putatively embedded word was somewhat truncated in the embedding word, or if the terminal element added to the putatively embedded word was a non-standard suffix of English. Investigation is required to find out how much truncation of embedded words speakers tolerate, and how much non-standard suffixation is likely to hinder preservation. Additionally, while semantic similarity has been assumed to be a requisite for a preservation relationship in this thesis, it seems plausible that there could be preservation from words which share no semantics with the embedding word, but which are phonologically similar (e.g.: *Apóllo* in *Apollinaris*, the latter being a brand of spring water; *Híspa*, a type of beetle, and *Hispaniola*, the island consisting of the Dominican Republic and Haiti). Existing psycholinguistic research may indicate where to begin in order to try and answer such questions. In turn, it will need to be considered whether fake cyclicity can handle every type of weak stress preservation relationship for which there is found to be evidence.

Finally, further attention must be given to the analysis of English strong preservation in the model of Stratal OT advocated in this thesis. It is clear from the foregoing chapters that there is much to be said and resolved with respect weak stress preservation in English, hence it has been the focus of this work. However, strong preservation is also not entirely without its controversy: notably, there are stratum two forms where strong preservation is expected, but weak preservation occurs instead, e.g. *nècessárily* (*nécessary*), *dòcuménted* (*dócument*) (Giegerich, 2005: §3.2.1.1). In addition to any empirical investigation, how we would want to deal with this in Stratal OT – notably, whether we would want to import the stratal overlap proposed by Giegerich (2005) – remains to be considered.

Appendices

Guide and key

(i) Words which exhibit stress preservation failure are italicised. Where preservation consistently fails, the syllable upon which secondary stress occurs in the embedding word is indicated (e.g. ‘init.’ indicates that secondary stress only falls on the initial syllable in the embedding word).

(ii) Words which have terminal elements not listed as standard suffixes in Marchand (1969) are indicated in bold for the appendices relating to left-edge stress preservation.

(iii) ‘(OED)’ indicates a putatively embedded word which was found in the *OED Online* and not in the pronouncing dictionary in question.

(iv) ‘Rem.’ is a word excluded from certain calculations in chapters 4 or 5 (details given in main text).

Appendix A: heavy-initial words from Wells (2000) where second-syllable preservation is expected

#HHL, #HLL and #HLH words

	Embedding word	Embedded word	
1.	**accèlerándo**	**accélerate**	**#HLL**
2.	accèlerátion	accélerate	#HLL
3.	**accèlerómeter**	**accélerate**	**#HLL**
4.	accèntuátion	accéntuate	#HHL
5.	accèptabílity	accéptable	#HHL
6.	accèssibílity	accéssible	Rem.
7.	admìnistrátion	admínistrate	Rem
8.	admìssibílity	admíssible	Rem.
9.	*anfractuosity*	*anfráctuous*	*#HHL*
10.	*antagonístic*	*antágonist*	*#HLL*
11.	antèriórity	antérior	#HHL
12.	*anticipation*	*antícipate*	*#HLL*
13.	*anticipatory*[1]	*antícipate*	*#HLL*
14.	***Antipodean***	***Antípodes***	*#HLL*
15.	artìculátory[2]	artículate	#HLL
16.	asphỳxiátion	asphýxiate	#HHL
17.	authènticátion	authénticate	#HHL
18.	authòritárian	authórity	#HLL
19.	**bactèriólogist**	**bactéria (bactèriólogy)**	**#HHL**
20.	**bactèriólogy**	**bactéria**	**#HHL**
21.	binòculárity	binócular	#HLL
22.	coàgulátion	coágulate	#HLL
23.	cohàbitátion	cohábit	#HLL
24.	combùstibílity	combústible	Rem.
25.	*concatenation*	*concátenate*	*#HLL*
26.	contèmporáneous	contémporary	#HHL
27.	coòperátion	coóperate	#HLL
28.	coòrdinátion	coórdinate	#HHL
29.	***Cù**nobelínus (init.)*	*Cun**ó**belin*	*#HLL/ #HHL Rem.*
30.	deàctivátion	deáctivate	#HHL
31.	decèlerátion	dècélerate	#HLL
32.	decònsecrátion	dècónsecrate	#HHL
33.	defènsibílity	defénsible	#HHL
34.	defìbrillátion	defíbrillate	Rem
35.	defòrestátion	defórest	Rem
36.	depòpulátion	depópulate	#HLL
37.	digèstibílity	digéstible	Rem.
38.	dimènsionálity	diménsional	#HHL
39.	disfòrestátion	disfórest	Rem

[1] Also *antícipatory*.
[2] Also *artículatory*.

40.	disìntegration	disíntegrate	#HHL
41.	**ejàculátio**	**ejáculate**	**#HLL**
42.	ejàculátion	ejáculate	#HLL
43.	elìcitátion	elícit	#HLL
44.	elùcidátion	elúcidate	#HHL
45.	elùcidátory	elúcidate	#HHL
46.	emàciátion	emáciate	#HHL
47.	emàsculátion	emásculate	Rem.
48.	encàpsulátion	encápsulate	#HHL
49.	envìronméntal	envíronment	#HHH/L
50.	eqùalitárian	eqúality	#HLL
51.	*(o)esophageal*	*(o)esóphagus*	*#HLL*
52.	eutròphicátion	eutróphic	#HLL
53.	evìscerátion	evíscerate	#HLL
54.	exàcerbátion	exácerbate	#HLL
55.	exàggerátion	exággerate	Rem.
56.	exàsperátion	exásperate	Rem.
57.	excìtabílity	excítable	#HHL
58.	excògitátion	excógitate	#HLL
59.	excòriátion	excóriate	#HL/HL Rem.
60.	exhìlarátion	exhílarate	#HLL
61.	exònerátion	exónerate	#HLL
62.	expànsibílity	expánsible	#HHL
63.	expàtriátion	expátriate	#HL/HL Rem.
64.	expèctorátion	expéctorate	#HHL
65.	expèriéntial	expérience	#HHL
66.	expòstulátion	expóstulate	Rem.
67.	exprèssionístic	expréssionist (OED)	Rem.
68.	expròpriátion	exprópriate	#HHL
69.	extèmporáneous	extémpore	#HHL
70.	extènsibílity	exténsible	#HHL
71.	extènsionálity	exténsional	#HHL
72.	extènuátion	exténuate	#HLL
73.	*exteriority*	*extérior*	*#HHL*
74.	extèrminátion	extérminate	#HHL
75.	extràpolátion	extrápolate	#HLL
76.	impàrtiálity	impártial	#HHL
77.	impèratíval	impérative	#HLL
78.	impèrsonátion	impérsonate	#HHL
79.	impètuósity	impétuous	#HLL
80.	implàusibility	impláusible	#HHL
81.	impòssibílity	impóssible	Rem.
82.	impràcticálity	impráctical	#HHL
83.	imprègnabílity	imprégnable	#HHL
84.	imprèssionístic	impréssionist	Rem.
85.	impròbabílity	impróbable	#HLL
86.	impròvabílity	impróvable	#HHL
87.	inàctivátion	ináctivate	#HHL

88.	inàudibílity	ináudible	#HHL
89.	incàpabílity	incápable	#HHL
90.	incàrcerátion	incárcerate	#HHL
91.	incìnerátion	incínerate	#HLL
92.	incònsequéntial	incónsequent	#HHL
93.	incòrporátion	incórporate	#HHL
94.	incrèdibílity	incrédible	#HLL
95.	incrìminátion	incríminate	#HLL
96.	incùrabílity	incúrable	#HHL
97.	inèdibílity	inédible	#HLL
98.	inèffabílity	inéffable	Rem.
99.	infàllibílity	infállible	Rem.
100.	infànticídal	infánticide	#HHL
101.	infàtuátion	infátuate	#HLL
102.	infèriórity	inférior	#HHL
103.	infìbulátion	infíbulate	#HLL
104.	*infinitival*	*infínitive*	*#HLL*
105.	inflèxibílity	infléxible	#HHL
106.	inqùisitórial	inqúisitor	#HLL
107.	inscrùtabílity	inscrútable	#HHL
108.	insèminátion	inséminate	#HLL
109.	insènsibílity	insénsible	#HHL
110.	insènsitívity	insénsitive	#HHL
111.	insìnuátion	insínuate	#HLL
112.	insòlubílity	insóluble	#HLL
113.	instàntiátion	instántiate	#HHL
114.	intàngibílity	intángible	#HHL
115.	**intèlligéntsia**	**intélligent**	**Rem.**
116.	intèrrogátion	intérrogate	Rem.
117.	intìmidátion	intímidate	#HLL
118.	intràctabílity	intráctable	#HHL
119.	intrànsitívity	intránsitive	#HHL
120.	intràvasátion	intrávasate	#HLL
121.	invàlidátion	inválidate	#HLL
122.	invèstigátion	invéstigate	Rem.
123.	invìgilátion	invígilate	#HLL
124.	invìgorátion	invígorate	#HLL
125.	invìncibílity	invíncible	#HHL
126.	invìsibílity	invísible	#HLL
127.	invòluntárily[3]	invóluntary	Rem
128.	*Nicomachean*	*Nicómachus*	*#HLL*
129.	*participation*	*partícipate*	*#HLL*
130.	perfèctibílity	perféctible	#HHL
131.	persèverátion	persévérate	#HLL
132.	potèntiálity	poténtial	#HHL
133.	potèntiátion	poténtiate	#HHL

[3] Also *invóluntarily*.

134.	prefàbricátion	prèfábricate	#HLL
135.	procràstination	procrástinate	Rem.
136.	*prognostication*	*prognósticate*	*Rem.*
137.	prolìferátion	prolíferate	#HLL
138.	propìtiátion	propítiate	#HLL
139.	*Pythagorean*	*Pythágoras*	*#HLL*
140.	redècorátion	redécorate	#HLL
141.	refòrestation	refórest	Rem
142.	revàluátion	reválue	#HLL
143.	*somnambulation*	*somnámbulate*	*#HHL*
144.	***Tèr**psichoréan (init.)*	*Terps**í**chore*	*#HLL*
145.	translìterátion	translíterate	#HLL
146.	transmìssibility	transmíssible	Rem.
147.	triàngulárity	triángular	#HHL
148.	triàngulation	triángulate	#HHL
149.	**tyrànnosáurus**	**tyránnosaur**	**Rem.**
150.	untòuchabílity	untóuchable	#HLL
151.	*Victoriana*	*Victórian*	*#HHL*
152.	vitùperátion	vitúperate	#HHL
153.	vocìferátion	vocíferate	#HLL

#H(=R)- words

154.	concìliátion	concíliate	#HLL
155.	condìtionálity	condítional	#HLL
156.	confèderátion	conféderate	#HLL
157.	congràtulátion	congrátulate	#HLL
158.	consìderátion	consider	#HLL
159.	conspìratórial	conspírator	#HLL
160.	contèmptibílity	contémptible	#HHL
161.	contìnuátion	contínue	#HLL
162.	convèntionálity	convéntional	#HHL
163.	convèrtibílity	convértible	#HHL
164.	convìviálity	convívial	#HLL

Appendix B: heavy-initial words from Jones (2003) where second-syllable preservation is expected

#HHL, #HLL and #HLH words

Embedding word	Embedded word	
1. **accèlerándo**	**accélerate**	**#HLL**
2. accèlerátion	accélerate	#HLL
3. accèntuátion	accéntuate	#HHL
4. accèptabílity	accéptable	#HHL
5. accèssibílity	accéssible	Rem.
6. admìssibílity	admíssible	Rem.
7. **ailùrophóbia**	**ailúrophobe**	**#HHL**
8. Amphìctyónic	Amphíctyon	*#HHL*
9. *anticipation*	*antícipate*	*#HLL*
10. *Antipodean*	*Antípodes*	*#HLL*
11. **artèriólar**	**artériole**	**#HHL**
12. artìculátion	artículate	#HLL
13. asphỳxiátion	asphýxiate	#HHL
14. authènticátion	authénticate	#HHL
15. *authoritarian*	*authórity*	*#HLL*
16. bactèriólogy	bactéria	#HHL
17. ***Bàr**tolomméo (init.)*	*Barth**ó**lomew*	*#HLL*
18. bisèxuálity	biséxual	#HHL
19. ***Boliviano***	***Bolívia***	***#HLL***
20. coàgulátion	coágulate	#HLL
21. confàbulátion	confábulate	#HLL
22. *cohabitation*	*cohábit*	*#HLL*
23. *concatenation*	*concátenate*	*#HLL*
24. *concelebration*	*concélebrate*	*#HLL*
25. *configuration*	*confígure*	*#HLL*
26. *conglomeration*	*conglómerate*	*#HLL*
27. conspìratórial	conspírator	#HLL
28. *contemporaneous*	*contémporary*	*#HHL*
29. ***cooperation***	***coóperate***	***#HLL***
30. coòrdinátion	coórdinate	#HHL
31. *d**è**consecrátion (init.)*	*dèc**ón**secrate*	*#HHL*
32. *de-escalation*	*dèéscalate*	*Rem.*
33. defìbrillátion	defíbrillate	Rem.
34. *defoliation*	*dèfóliate*	*#HHL*
35. *deforestation*	*dèfórest*	*Rem.*
36. *depopulátion*	*dèpópulate*	*#HLL*
37. *deracination*	*derácinate*	*#HLL*
38. *d**è**regulátion (init.)*	*dèr**é**gulate*	*#HLL*
39. *desalination*	*dèsálinate*	*#HLL*
40. *desegregation*	*dèségregate*	*#HLL*
41. digèstibílity	digéstible	Rem.
42. dilàtabílity	dilátable (OED)	#HHL

43. *dissatisfaction*	*dissátisfy*	*Rem.*
44. *dissatisfactory*	*dissátisfy*	*Rem.*
45. ejàculátion	ejáculate	#HLL
46. elìcitátion	elícit	#HLL
47. emàciátion	emáciate	#HHL
48. emàsculation	emásculate	*Rem.*
49. encàpsulátion	encápsulate	#HHL
50. envìronméntal	envíronment	#HHH/L
51. *(o)esophageal*	*esóphagus*	*#HLL*
52. evìscerátion	evíscerate	#HLL
53. exàcerbátion	exácerbate	#HLL
54. exàggerátion	exággerate	Rem.
55. exàminátion	exámine	#HLL
56. exàsperátion	exásperate	Rem.
57. exchàngeabílity	exchángeable	#HHL
58. excìtabílity	excítable	#HHL
59. exclùsivístic	exclúsive	#HHL
60. excòriátion	excóriate	#HL/HL Rem.
61. excrùciátion	excrúciate	#HHL
62. *exfoliation*	*exfóliate*	*#HHL*
63. exhìlarátion	exhílarate	#HLL
64. exònerátion	exónerate	#HLL
65. expànsibílity	expánsible	#HHL
66. expàtiátion	expátiate	#HHL
67. expèctorátion	expéctorate	#HHL
68. expèriéntial	expérience	#HHL
69. *experimental*	*expériment*	*#HLL*
70. expòstulátion	expóstulate	Rem.
71. exprèssionístic	expréssionist (OED)	Rem.
72. *expropriation*	*exprópriate*	*#HHL*
73. *extemporaneous*	*extémpore, extémporary*	*#HHL*
74. extènsibílity	exténsible	#HHL
75. extènuátion	exténuate	#HLL
76. extèrminátion	extérminate	#HHL
77. extràpolátion	extrápolate	#HLL
78. exùviátion	exúviate	#HHL
79. *foreseeability*	*foreséeable*	*#HHL*
80. hermàphrodític	hermáphrodite	#HLL
81. *humiliation*	*humíliate*	*#HLL*
82. hypòthecátion	hypóthecate	#HLL
83. idèntifíable[4]	idéntify	#HHL
84. *impartiality*	*impártial*	*#HHL*
85. *impassability*	*impássable*	*Rem.*
86. *impeccability*	*impéccable*	*Rem.*
87. ***impedimenta***	***impédiment***	***#HLL***
88. impèrsonálity	impérsonal	#HHL

[4] Also *idéntifiable*.

89. impèrsonátion	impérsonate	#HHL
90. implàcabílity	implácable	#HLL
91. *implausibility*	*impláusible*	*#HHL*
92. *impossibility*	*impóssible*	*Rem.*
93. *impracticality*	*impráctical*	*#HHL*
94. impregnability	imprégnable	#HHL
95. imprèssibílity	impréssible	Rem.
96. *improbability*	*impróbable*	*#HLL*
97. *impropriation*	*imprópriate*	*#HHL*
98. impròvabílity	impróvable	#HHL
99. *imputability*	*impútable* (OED)	#HHL
100. incàpabílity	incápable	#HHL
101. *incarcerátion*	*incárcerate*	*#HHL*
102. incìnerátion	incínerate	#HLL
103. incònseqúential	incónsequent	#HHL
104. incòrporátion	incórporate	#HHL
105. incrèdibílity	incrédible	#HLL
106. incrìminátion	incríminate	#HLL
107. incùrabílity	incúrable	#HHL
108. indèlibílity	indélible	#HLL
109. indòctrinátion	indóctrinate	#HHL
110. infàllibílity	infállible	Rem.
111. infàtuátion	infátuate	#HLL
112. infibulátion	infíbulate	#HLL
113. infràngibílity	infrángible	#HHL
114. inhàbitátion	inhábit	#HLL
115. *inquisitorial*	*inqúisitor*	*#HLL*
116. insàtiabílity	insátiable	#HHL
117. inscrùtabílity	inscrútable	#HHL
118. insèminátion	inséminate	#HLL
119. insènsibílity	inséensible	#HHL
120. insènsitívity	inséensitive	#HHL
121. insìnuátion	insínuate	#HLL
122. insòlubílity	insóluble	#HLL
123. intàngibílity	intángible	#HHL
124. **intèlligéntsia**	**intélligent**	**Rem.**
125. intèrcalátion	intércalate	#HHL
126. intèrpolátion	intérpolate	#HHL
127. intèrpretátion	intérpret	#HHL
128. intèrrogátion	intérrogate	Rem.
129. intìmidátion	intímidate	#HLL
130. intòxicátion	intóxicate	#HHL
131. intràctabílity	intráctable	#HHL
132. intrànsitívity	intránsitive	#HHL
133. invàlidátion	inválidate	#HLL
134. invèstigátion	invéstigate	Rem.
135. invìgilátion	invígilate	#HLL
136. invìncibílity	invíncible	#HHL

137.	invìgorátion	invígorate	#HLL
138.	invìsibílity	invísible	#HLL
139.	*municipality*	*municipal*	*#HLL*
140.	*Nicomachean*	*Nichómachus*	*#HLL*
141.	*participation*	*partícipate*	*#HLL*
142.	perfèctibílity	perféctible	#HHL
143.	perpètuátion	perpétuate	#HLL
144.	persèverátion	perséverate	#HLL
145.	potèntiálity	poténtial	*#HHL*
146.	***prè**destinátion (init.)*	*prèd**é**stine*	*Rem.*
147.	*prefabrication*	*prè-fábricate*	*#HLL*
148.	*premeditation*	*prèméditate*	*#HLL*
149.	*preoccupation*	*preóccupy*	*Rem.*
150.	prepònderátion	prepónderate	#HHL
151.	procràstinátion	procrástinate	Rem.
152.	prognòsticátion	prognósticate	Rem.
153.	prolìferátion	prolíferate	#HLL
154.	**pronùnciaménto**	**pronóunce**	**#HHL**
155.	*Pythagorean*	*Pythágoras*	*#HLL*
156.	***rè**activátion (init)*	*rè**ác**tivate*	*#HHL*
157.	***rè**generátion (init.)*	*règ**é**nerate*	*#HLL*
158.	reòccupátion	reóccupy	Rem.
159.	***rè**patriátion (init.)*	*rèp**á**triate*	*#HLL/HHL Rem.*
160.	Shakespèreána	Shakespérean	#HHL
161.	***Tèrp**sichoréan (init.)*	*Terps**í**chore*	*#HLL*
162.	*transliteration*	*translíterate*	*#HLL*
163.	*transmissibility*	*transmíssible*	*Rem.*
164.	*transmutability*	*transmútable (OED)*	*#HHL*
165.	**tyrànnosáurus**	**tyránnosaur**	**Rem.**
166.	unchàngeabílity	unchángeable	#HHL
167.	unpòpulárity	unpópular	#HLL
168.	unpràcticálity	unpráctical	#HHL
169.	unpùnctúality	unpúnctual	#HHL
170.	*unsuitability*	*unsúitable*	*#HHL*
171.	untòuchabílity	untóuchable	#HLL
172.	*Victoriana*	*Victórian*	*#HHL*
173.	vitùperátion	vitúperate	#HHL
174.	vocìferátion	vocíferate	#HLL

#H(=R)- words

175.	advìsabílity	advísable	#HHL
176.	combùstibílity	combústible	Rem.
177.	compàtibílity	compátible	#HLL
178.	comprèssibílity	compréssible	Rem.
179.	concìliátion	concíliate	#HLL
180.	condùctibílity	condúctible	#HHL
181.	confèderátion	conféderate	#HLL
182.	confòrmabílity	confórmable	#HHL

183.	congèniálity	congénial	#HHL
184.	congràtulátion	congrátulate	#HLL
185.	conjùnctivítis	conjúnctive	#HHL
186.	connùbuálity	connúbial	#HHL
187.	consìderátion	consíder(ate)	#HLL
188.	consòlidátion	consólidate	#HLL
189.	contàminátion	contáminate	#HLL
190.	contèmptibílity	contémptible	#HHL
191.	contìnuátion	contínue	#HLL
192.	contràctibílity	contráctible	#HHL
193.	convèntionálity	convéntional	#HHL
194.	convèrtibílity	convértible	#HHL
195.	convìviálity	convívial	#HLL
196.	subsìdiárity	subsídiary	#HLL
197.	substàntiálity	substántial	#HHL
198.	substàntiátion	substántiate	#HHL

Appendix C: heavy-initial words from Wells (2000) where initial-syllable preservation is expected

#HLL words

Embedding word	Embedded word	
1. àlienátion	álienate	
2. àmiabílity	ámiable	
3. àmplificátion	ámplify	
4. àxiomátic	áxiom	
5. bèautificátion	béautify	
6. càlcificátion	cálcify	
7. càlculabílity	cálculable	
8. cèntrifugátion	céntrifuge	
9. còdificátion	códify	
10. ètiolátion	étiolate	
11. fòrtificátion	fórtify	
12. frùctificátion	frúctify	
13. gèntrificátion	géntrify	
14. glòrificátion	glórify	
15. ìnfinitésimal	ínfinite	
16. lìgnificátion	lígnify	
17. màgnificátion	mágnify	
18. màrketabílity	márketable	
19. mèliorátion	méliorate	
20. mòrtificátion	mórtify	
21. mùltiplicátion	múltiply	
22. nòtificátion	nótify	
23. pàncreatítis	páncreas	
24. pèrmeabílity	pérmeable	
25. *phantasmagoria*	*phántàsm*	*Rem.*
26. pràcticabílity	prácticable	
27. pùrificátion	púrify	
28. pùrificátory[5]	púrify	
29. qùantificátion	qúantify	
30. rèctificátion	réctify	
31. rèificátion	réify	
32. sànctificátion	sánctify	
33. scàrificátion	scárify	
34. sìgnificátion	sígnify	
35. sìmplificátion	símplify	
36. spèechificátion	spéechify	
37. stùltificátion	stúltify	
38. ùnificátion	únify	
39. vàriabílity	váriable	
40. vàriegátion	váriegate (OED)	

[5] Also *púrificàtory*.

41. vèrsificátion	vérsify

#HL(=R)L words

50. ànswerabílity	ánswerable
51. dìatomáceous	díatom
52. jùxtaposítion	júxtapose[6]
53. mènsurabílity	ménsurable
54. Pàntagruélian	Pántagruel
55. vùlnerabílity	vúlnerable

Sometimes or always #HLH or #HHH

48. ànglicizátion	ánglicize
49. àrgumentátion	árgument
50. àrmamentárium	**ármament**
51. bòwdlerizátion	bówdlerize
52. brùtalizátion	brútalize
53. càrbonizátion	cárbonize
54. càuterizátion	cáuterize
55. cèntralizátion	céntralize
56. chàptalizátion	cháptalize
57. Chrìstianizátion	Chrístianize
58. crèolizátion	créolize
59. dìeselizátion	díeselize (OED)
60. dràmatizátion	drámatize
61. èqualizátion	équalize
62. èndocrinólogy	**éndocrine**
63. fèrtilizátion	fértilize
64. Fìnlandizátion	Fínlandize (OED)
65. fòrmalizátion	fórmalize
66. gàlvanizátion	gálvanize
67. glòbalizátion	glóbalize (OED)
68. hàrmonizátion	hármonize
69. hùmanizátion	húmanize
70. hỳbridizátion	hýbridize
71. ìdeográphic	ídeograph
72. ìdolizátion	ídolize
73. ìmplementátion	ímplement
74. ìmprovisátion	ímprovise
75. ìnstrumentátion	ínstrument
76. ìodizátion	íodize
77. ìonizátion	íonize
78. làicizátion	láicize
79. lègalizátion	légalize
80. lìonizátion	líonize
81. lòcalizátion	lócalize
82. màgnetizátion	mágnetize

[6] *Jùxtapóse* is also given by Wells (2000).

83.	màximizátion	máximize
84.	mòbilizátion	móbilize
85.	mòngrelizátion	móngrelize
86.	nàturopáthic	náturopath
87.	nàsalizátion	násalize
88.	nèutralizátion	néutralize
89.	nòrmalizátion	nórmalize
90.	nòtarizátion	nótarize
91.	òptimizátion	óptimize
92.	òrganizátion	órganize
93.	òrientátion	órientate
94.	òrnamentátion	órnament
95.	òxidizátion	óxidize
96.	pàsteurizátion	pásteurize
97.	pàuperizátion	páuperize
98.	pènalizátion	pénalize
99.	plèsiosáurus	plésiosaur
100.	plùralizátion	plúralize
101.	pòlarizátion	pólarize
102.	prìvatizátion	prívatize
103.	pùlverizátion	púlverize
104.	rhòtacizátion	rhótacize
105.	ròmanizátion	rómanizátion
106.	schèmatizátion	schématize
107.	sènsitizátion	sénsitize
108.	sòcializátion	sócialize
109.	spìrantizátion	spírantize
110.	stàbilizátion	stábilize
111.	stàndardizátion	stándardize
112.	stìgmatizátion	stígmatize
113.	sùbsidizátion	súbsidize
114.	sùlcalizátion	súlcalize
115.	sỳnchronizátion	sýnchronize
116.	tràumatizátion	tráumatize
117.	ùrbanizátion	úrbanize
118.	ùtilizátion	útilize
119.	vàporizátion	váporize
120.	vèlarizátion	vélarize
121.	vèrbalizátion	vérbalize
122.	vèrnalizátion	vérnalize
123.	vìctimizátion	víctimize
124.	vòcalizátion	vócalize
125.	vùlcanizátion	vúlcanize
126.	vùlgarizátion	vúlgarize
127.	wìnterizátion	wínterize

Appendix D: heavy-initial words from Jones (2003) where initial-syllable preservation is expected

#HLL words

Embedding word	Embedded word	
1. àlienátion	álienate	
2. àmiabílity	ámiable	
3. àmplificátion	ámplify	
4. àmplificátory	ámplify	
5. àxiomátic	áxiom	
6. bèautificátion	béautify	
7. càlcificátion	cálcify	
8. còdificátion	códify	
9. còmbinatórial	cómbinatory[7]	
10. comparability	*cómparable*	*Rem.*
11. fàlsificátion	fálsify	
12. fòrtificátion	fórtify	
13. frùctificátion	frúctify	
14. gèntrificátion	géntrify	
15. glòrificátion	glórify	
16. ìnfinitésimal	ínfinite	
17. màgnificátion	mágnify	
18. màrketabílity	márketable	
19. mèdiatórial	médiator	
20. mèliorátion	méliorate	
21. mòrtificátion	mórtify	
22. mùltiplicátion	múltiply	
23. nòtificátion	nótify	
24. pèrmeabílity	pérmeable	
25. phantasmagoria	*phántàsm*	*Rem.*
26. pràcticabílity	prácticable	
27. pùrificátion	púrify	
28. pùrificátory[8]	púrify	
29. qùantificátion	qúantify	
30. ràreficátion	rárefy	
31. rèctificátion	réctify	
32. rèificátion	réify	
33. sànctificátion	sánctify	
34. sìgnificátion	sígnify	
35. sìmplificátion	símplify	
36. spèechificátion	spéechify	
37. stùltificátion	stúltify	
38. ùnificátion	únify	
39. vàriabílity	váriable	
40. vèrsificátion	vérsify	

[7] Or *còmbinátory*.
[8] Also *púrificàtory*.

41. vìndicabílity	víndicable	

#HHL words

42. àlbuminúria	álbumin	

#HL(=R)L words

43. ànswerabílity	ánswerable	
44. dìatomáceous	díatom	
45. jùxtaposítion	júxtapose	
46. mènsurabílity	ménsurable	
47. pònderabílity	pónderable	
48. vùlnerabílity	vúlnerable	

Sometimes or always #HLH or #HHH words

49. àlkalizátion	álkalize	
50. ànglicizátion	ánglicize	
51. ànticlimáctic	ánticlìmax[9]	
52. àrgumentátion	árgument	
53. Bàlkanizátion	Bálkanize	
54. bòwdlerizátion	bówdlerize	
55. càrbonizátion	cárbonize	
56. càuterizátion	cauterize	
57. cèntralizátion	céntralize	
58. dìaphragmátic	díaphragm	Rem.
59. dràmatizátion	drámatize	
60. èqualizátion	équalize	
61. fàctorizátion	fáctorize	
62. fèrtilizátion	fértilize	
63. Fìnlandizátion	Fínlandize	Rem.
64. flùidizátion	flúidize	
65. fòrmalizátion	fórmalize	
66. gèrmanizátion	gérmanize	
67. glòbalizátion	glóbalize	
68. hàrmonizátion	hármonize	
69. hùmanizátion	húmanize	
70. hỳbridizátion	hýbridize	
71. hỳpnotizátion	hýpnotize	
72. ìdolizátion	ídolize	
73. ìmplementátion	ímplement	
74. ìnstrumentátion	ínstrument	
75. ìonizátion	íonize	
76. lègalizátion	légalize	
77. lìonizátion	líonize	
78. lòcalizátion	lócalize	
79. màximizátion	máximize	
80. mòbilizátion	móbilize	

[9] Also *ànticlímax*.

81.	nàsalizátion	nasalize
82.	nàturopáthic	náturopath
83.	nèutralizátion	néutralize
84.	nòrmalizátion	nórmalize
85.	òptimizátion	óptimize
86.	òrganizátion	organize
87.	òrientátion	órientate
88.	òxidizátion	óxidize
89.	pàrliamentárian	párliament
90.	pàssivizátion	pássivize
91.	pàsteurizátion	pásteurize
92.	pàuperizátion	páuperize
93.	prìvatizátion	prívatize
94.	pùlverizátion	púlverize
95.	ràndomizátion	rándomize
96.	ròmanizátion	rómanize
97.	sènsitizátion	sénsitize
98.	stàbilizátion	stábilize
99.	stàndardizátion	stándardize
100.	stìgmatizátion	stígmatize
101.	sùlcalizátion	súlcalize
102.	sỳnchronizátion	sýnchronize
103.	tàntalizátion	tántalize
104.	tèrgiversátion	térgiversate
105.	tèutonizátion	téutonize
106.	trànquilizátion	tránquilize
107.	tràumatizátion	tráumatize
108.	ùrbanizátion	úrbanize
109.	ùtilizátion	útilize
110.	vàporizátion	váporize
111.	vèlarizátion	vélarize
112.	vèrbalizátion	vérbalize
113.	vìctimizátion	víctimize
114.	vòcalizátion	vócalize
115.	vùlcanizátion	vúlcanize
116.	vùlgarizátion	vúlgarize

Appendix E: #LLL and #LLH words from Wells (2000) where second-syllable preservation is expected

#LLL and #LLH words

Embedding word	Embedded word		
1. affòrestátion	affórest	Rem.	#LLL
2. anastomosis	*anástomose*	*Rem.*	*#LLL*
3. assèverátion	asséverate	Rem.	#LLL
4. assìbilátion	assíbilate	Rem.	#LLL
5. binòculárity	binócular		#LLL
6. bolìviáno	**Bolívia**		**#LLL**
7. debìlitátion	debílitate		#LLL
8. decàpitátion	decápitate		#LLL
9. defòrestátion	defórest	Rem.	#LLL
10. degènerátion	degénerate		#LLL
11. delìberátion	delíberate		#LLL
12. delìmitátion	delímitate		#LLL
13. delìneátion	delíneate		#LLL
14. demòdulátion	demódulate		#LLL
15. denòminátion	denóminate		#LLL
16. desìderáta	**desíderate (OED)**		**#LLL**
17. devèlopméntal	devélopment	Rem.	#LLH
18. dilàpidátion	dilápidate		#LLL
19. dimìnuéndo	**dimínish**		**#LLL**
20. discrìminátion	discríminate	Rem.	#LLL
21. dissatisfaction	*dissátisfy*	*Rem.*	*#LLL*
22. dissèminátion	disséminate	Rem.	#LLL
23. dissimilarity	*dissimilar*	*Rem.*	*#LLL*
24. dissimilation	*dissimilate*	*Rem.*	*#LLL*
25. dissìmulátion	dissímulate	Rem.	#LLL
26. dissòlubílity	dissóluble	Rem.	#LLL
27. divìsibílity	divísible		#LLL
28. ecònométric	**ecónomy**		**#LLL**
29. ejàculátio	**ejáculate**		**#LLL**
30. ejàculátion	ejáculate		#LLL
31. elàborátion	eláborate		#LLL
32. elìcitátion	elícit		#LLL
33. elìminátion	elíminate		#LLL
34. Elìzabéthan	Elízabeth		#LLL
35. ephèmerálity	ephémeral		#LLL
36. epìscopálian	epíscopal	Rem.	#LLL
37. eqùalitárian	eqúality		#LLL
38. eqùivocátion	eqúivocate		#LLL
39. eràdicátion	erádicate		#LLL
40. evàluátion	eváluate		#LLL
41. evàporátion	eváporate		#LLL
42. evìscerátion	evíscerate		#LLL

43. felìcitátion	felícitate		#LLL
44. gestìculátion	gestículate	Rem.	#LLL
45. habìlitátion	habílitate		#LLL
46. habìtuátion	habítuate		#LLL
47. herèditárily[10]	heréditary		#LLL
48. Hispaniola	*Hispánic*	*Rem.*	*#LLL*
49. hòrripilátion	*horripilate (OED)*	*Rem.*	*#LLL*
50. illègibílity	illégible	Rem.	#LLL
51. imàginátion	imágine		#LLL
52. inamorato	*inámorate (OED)*	*Rem.*	*#LLL*
53. inìtiátion	inítiate		#LLL
54. inòculátion	inóculate	Rem.	#LLL
55. iràscibílity	iráscible		#LLL
56. irràtionálity	irrátional	Rem.	#LLL
57. irregularity	*irregular*	*Rem.*	*#LLL*
58. irrèsolútion	irrésolute	Rem.	#LLL
59. legìtimátion	legítimate		#LLL
60. miscegenation	*miscégenate* (OED)		*#LLL*
61. misògynístic	misógynist		#LLL
62. oblìterátion	oblíterate		#LLL
63. orìginálity	oríginal		#LLL
64. orìginátion	oríginate		#LLL
65. perìpherálity	perípheral		#LLL
66. phenòmenólogy	**phenómenon**		**#LLL**
67. precìpitátion	precípitate		#LLL
68. predèstinátion	predéstinate	Rem.	#LLL
69. premèditátion	preméditate		#LLL
70. preòccupátion	preóccupy	Rem.	#LLL
71. prevàricátion	prevaricate		#LLL
72. recìprocálity	recíprocal		#LLL
73. recìprocátion	recíprocate		#LLL
74. recrìminátion	recríminate		#LLL
75. reèducátion	reéducate		#LLL
76. refrìgerátion	refrígerate		#LLL
77. regènerátion	regénerate		#LLL
78. reìterátion	reíterate		#LLL
79. relìgiósity	relígious		#LLL
80. repàtriátion	repátriate		#LLL
81. resùscitátion	resúscitate		#LLL
82. retàliátion	retáliate		#LLL
83. retìculátion	retículate		#LLL
84. syllàbicátion	syllábicate	Rem.	#LLL
85. theàtricálity	theátrical		#LLL
86. tyrànnosáurus	**tyránnosaur**	Rem.	#LLL
87. vaticination	*vatícinate*		*#LLL*
88. verìdicálity	verídical		#LLL

[10] Also *heréditarily*.

#L(=R)LL and #LLH words

89. accòmmodátion	accómmodate	#LLL
90. accrèditátion	accrédit	#LLL
91. agglòmerátion	agglómerate	#LLL
92. allìterátion	allíterate	#LLL
93. anòmalístic	anómaly	#LLL
94. apòcalýptic	apócalypse	#LLL
95. **Apòllodórus**	**Apóllo**	#LLL
96. apòlogétic	apólogy	#LLL
97. assìmilátion	assímilate	#LLL
98. Assỳriólogy	**Assýria**	#LLL
99. attènuátion	atténuate	#LLL
100. capìtulátion	capítulate	#LLL
101. caprìcciós o	**capríccio**	#LLL
102. collàborátion	colláborate	#LLL
103. commèmorátion	commémorate	#LLL
104. commìserátion	commíserate	#LLL
105. corròberátion	corróberate	#LLL
106. domèsticátion	doméstic ate	#LLL
107. facìlitátion	facílitate	#LLL
108. famìliárity	famíliar	#LLL
109. grammàticálity	grammátical	#LLL
110. machìcolátion	machícolate	#LLL/H
111. majòritárian[11]	majority	#LLL
112. manìpulátion	manípulate	#LLL
113. matrìculátion	matrículate	#LLL
114. monòpolístic	monópolist	#LLL
115. offìciátion	offíciate	#LLL
116. partìculárity	partícular	#LLL
117. perpètuátion	perpétuate	#LLL
118. solìcitátion	solícit	#LLL
119. sophìsticátion	sophísticate	#LLL
120. suggèstibílity	suggéstible	#LLL
121. tachìstoscópic	tachístoscope	#LLL

[11] The terminal element *-arian* can attach to bound roots – *agrárian* – and so there need be no intermediary form between *majoritarian* and *majority*.

Appendix F: #LLL and #LLH words from Jones (2003) where second-syllable preservation is expected

#LLL and #LLH words

	Embedding word	Embedded word		
1.	allìterátion	allíterate	Rem.	#LLL
2.	***à**nastomósis (init.)*	*an**á**stomose (OED)*	Rem.	#LLL
3.	assèverátion	assévérate	Rem.	#LLL
4.	assìbilátion	assíbilate	Rem.	#LLL
5.	***Boliviano***	***Bolívia***		***#LLL***
6.	collàborátion	colláborate	Rem.	#LLL
7.	commìserátion	commíserate	Rem.	#LLL
8.	deàminátion	deáminate		#LLL
9.	debìlitátion	debílitate		#LLL
10.	decàpitátion	decápitate		#LLL
11.	decrèpitátion	decrépit		#LLL
12.	defòrestátion	defórest	Rem.	#LLL
13.	degènerátion	degénerate		#LLL
14.	delìberátion	delíberate		#LLL
15.	delìmitátion	delímit		#LLL
16.	delìneátion	delíneate		#LLL
17.	demòdulátion	demódulate		#LLL
18.	denòminátion	denóminate		#LLL
19.	desìderátion	desíderate		#LLL
20.	**desìderátum**	**desíderate**		**#LLL**
21.	devàluátion	deválue		#LLL
22.	devèlopméntal	development	Rem.	#LLH
23.	digèstibílity	digéstible	Rem.	#LLL
24.	dilàpidátion	dilápidate		#LLL
25.	**dimìnuéndo**	**dimínish**		**#LLL**
26.	discrìminátion	discríminate	Rem.	#LLL
27.	*discolouration*	*discólour*	*Rem.*	*#LLL*
28.	dissèminátion	disséminate	Rem.	#LLL
29.	dissìmilárity	dissímilar	Rem.	#LLL
30.	dissìmilátion	dissímilate	Rem.	#LLL
31.	dissìmulátion	dissímulate	Rem.	#LLL
32.	dissòlubílity	dissóluble	Rem.	#LLL
33.	divìsibílity	divísible		#LLL
34.	elàborátion	eláborate		#LLL
35.	elìminátion	elíminate		#LLL
36.	Elìzabéthan	Elízabeth		#LLL
37.	Elìzabéthian	Elízabeth		#LLL
38.	emàsculátion	emásculate	Rem.	#LLL
39.	ephèmerálity	ephémeral		#LLL
40.	epìscopálian	epíscopal	Rem.	#LLL
41.	eqùivocátion	eqúivocate		#LLL
42.	eràdicátion	erádicate		#LLL

43.	(o)esòphagéal	(o)esóphagus		#LLL
44.	evàcuátion	evácuate		#LLL
45.	evàluátion	eváluate		#LLL
46.	evìscerátion	evíscerate		#LLL
47.	felìcitátion	felícitate		#LLL
48.	gestìculátion	gestículate	Rem.	#LLL
49.	habìlitátion	habílitate		#LLL
50.	*Hispaniola*	*Hispánic*	*Rem.*	*#LLL*
51.	*horripilation*	*horrípilate (OED)*	*Rem.*	*#LLL*
52.	*illegibility*	*illégible*	*Rem.*	*#LLL*
53.	illìberálity	illíberal	Rem.	#LLL
54.	illògicálity	illógical	Rem.	#LLL
55.	imàginátion	imágine		#LLL
56.	immìscibility	immíscible	Rem.	#LLL
57.	immòderátion	immóderate	Rem.	#LLL
58.	*inamorato*	*inámorate (OED)*	*Rem.*	*#LLL*
59.	inòculátion	inóculate	Rem.	#LLL
60.	inòsculátion	inósculate	Rem.	#LLL
61.	iràscibílity	iráscible		#LLL
62.	irràtionálity	irrátional	Rem.	#LLL
63.	irrègulárity	irrégular	Rem.	#LLL
64.	irrèsolútion	irrésolute	Rem.	#LLL
65.	legìtimátion	legítimate		#LLL
66.	machìcolátion	machícolate	Rem.	#LLL/H
67.	*mìscegenation (init.)*	*miscégenate (OED)*		*#LLL*
68.	orìginálity	oríginal		#LLL
69.	orìginátion	oríginate		#LLL
70.	**phenòmenólogy**	**phenómenon**		**#LLL**
71.	precìpitátion	precípitate		#LLL
72.	predèstinátion	predéstinate	Rem.	#LLL
73.	predòminátion	predóminate		#LLL
74.	premèditátion	preméditate		#LLL
75.	preòccupátion	preóccupy	Rem.	#LLL
76.	prevàricátion	preváricate		#LLL
77.	reàllocátion	reállocate	Rem.	#LLL
78.	recìprocátion	recíprocate		#LLL
79.	refrìgerátion	refrígerate		#LLL
80.	regènerátion	regénerate		#LLL
81.	relìgiósity	relígious		#LLL
82.	reòccupátion	reóccupy	Rem.	#LLL
83.	repàtriátion	repátriate		#LLL
84.	resùscitátion	resúscitate		#LLL
85.	retàliátion	retáliate		#LLL
86.	retìculátion	retículate		#LLL
87.	syllàbicátion	syllábicate	Rem.	#LLL
88.	thèatricálity	theátrical		#LLL
89.	**tyrànnosáurus**	**tyránnosaur**	Rem.	#LLL
90.	*vaticination*	*vatícinate*		*#LLL*

#L(=R)LL and #L(=R)LH words

91.	abòminátion	abóminate	#LLL
92.	accòmmodátion	accómmodate	#LLL
93.	accrèditátion	accrédit	#LLL
94.	acètylátion	acétylate	#LLL
95.	agglòmerátion	agglómerate	#LLL
96.	**Amèricána**	**Américan**	**#LLL**
97.	anàchronístic	anáchronous	#LLL
98.	apòcalyptic	apócalypse	#LLL
99.	**Apòllodórus**	**Apóllo**	#LLL
100.	apòlogétic	apólogy	#LLL
101.	assàssinátion	assássinate	#LLL
102.	assìmilátion	assímilate	#LLL
103.	**Assỳriólogy**	**Assýria**	#LLL
104.	attènuátion	atténuate	#LLL
105.	capìtulátion	capítulate	#LLL
106.	**caprìccioso**	**capríccio**	#LLL
107.	commèmorátion	commémorate	#LLL
108.	domèsticátion	doméstícate	#LLL
109.	facìlitátion	facílitate	#LLL
110.	famìliárity	famíliar	#LLL
111.	grammàticálity	grammátical	#LLL
112.	manìpulátion	manípulate	#LLL
113.	matrìculátion	matriculate	#LLL
114.	monòpolístic	monópolist	#LLL
115.	oblìterátion	oblíterate	#LLL
116.	partìculárity	partícular	#LLL
117.	propìtiátion	propítiate	#LLL
118.	solìcitátion	solícit	#LLL
119.	sophìsticátion	sophísticate	#LLL
120.	suggèstibílity	suggéstible	#LLL
121.	surrèalistic	surréalist (OED)	#LLL

Appendix G: #LHL and #LHH words from Wells (2000)

#LHL and #LHH words

	Embedding word	Embedded word	
1.	accùlturátion	accúlturate	#LHL
2.	**aràchnophóbia**	**aráchnid (OED)**	**#LHL/H**
3.	behàviourístic	beháviour	#LHL
4.	connùbiálity	connúbial	#LHL
5.	defènsibílity	defénsible	#LHL
6.	degràdabílity	degrádable	#LHL
7.	depèndabílity	depéndable	#LHL
8.	deprèciátion	depréciate	#LHL
9.	desìrabílity	desírable	#LHL
10.	despòliátion	despóil	#LHL
11.	detòxicátion	detóxicate	#LHL
12.	dirèctionálity	diréctional	#LHL
13.	dispòsabílity	dispósable	#LHL
14.	dissòciátion	dissóciate	#LHL
15.	**divèrtiménto**	**divért**	**#LHL**
16.	**Ecclèsiástes**	**Ecclésiast (OED)**	**#LHL**
17.	ecclèsiástic	Ecclésiast (OED)	#LHL
18.	effèctuátion	efféctuate	#LHL
19.	elèctrocútion	eléctrocute	#LHL
20.	emàncipátion	emáncipate	#LHL
21.	enùmerátion	enúmerate	#LHL
22.	enùnciátion	enúnciate	#LHL
23.	essèntiálity	esséntial	#LHL
24.	evàngelístic	evángelist	#LHL
25.	evèntuálity	evéntual	#LHL
26.	illùminátion	illúminate	#LHL
27.	immùtabílity	immútable	#LHL
28.	inèbriátion	inébriate	#LHL
29.	*inauguration*	*ináugurate*	*#LHL*
30.	irràdiátion	irrádiate	#LHL
31.	***listeriosis***	***listéria***	***#LHL***
32.	Napòleonic	Napóleon	#LHL
33.	negòtiabílity	negótiable	#LHL
34.	negòtiátion	negótiate	#LHL
35.	pecùliárity	pecúliar	#LHL
36.	peràmbulátion	perámbulate	#LHL
37.	*posteriority*	*postérior*	*#LHL*
38.	predìctabílity	predíctable	#LHL
39.	prevèntabílity	prevéntable	#LHL
40.	reàctivátion	reáctivate	#LHL
41.	recùperátion	recúperate	#LHL
42.	redùcibílity	redúcible	#LHL
43.	redùplicátion	redúplicate	#LHL

44.	regùrgitátion	regúrgitate	#LHL
45.	relìabílity	relíable	#LHL
46.	rejùvenátion	rejúvenate	#LHL
47.	remèdiátion	remédiate (OED)	#LHL
48.	remòvabílity	remóvable	#LHL
49.	remùnerátion	remúnerate	#LHL
50.	renèwabílity	renéwable	#LHL
51.	renùnciátion	renóunce	#LHL
52.	repùdiátion	repúdiate	#LHL
53.	resòlvabílity	resólvable	#LHL
54.	respèctabílity	respéctable	#LHL
55.	respònsibílity	respónsible	#LHL
56.	retrìevabílity	retríevable	#LHL
57.	revèrberátion	revérberate	#LHL
58.	revèrsibílity	revérsible	#LHL
59.	**sotèriólogy**	**sotérial**	**#LHL**
60.	supèriórity	supérior	#LHL
61.	**tubèrculósis**	**tubérculous**	**#LHL**
62.	venèreólogy	venéreal	#LHL

#L(=R)HL and #LHH words

63. abbrèviátion	abbréviate	#LHL
64. abbrèviátory[12]	abbréviate	#LHL
65. accòuntabílity	accóuntable	#LHL
66. accùsatíval	accúsative	#LHL
67. accùsatórial	accúse	#LHL
68. adàptabílity	adáptable	#LHL
69. adjùdicátion	adjúdicate	#LHL
70. adùlterátion	adúlterate	#LHL
71. advìsabílity	advísable	#LHL
72. affòrdabílity	affórdable	#LHL
73. allèviátion	alléviate	#LHL
74. amènabílity	aménable	#LHL
75. **appèndicéctomy**	**appéndix**	**#LHL**
76. **appèndicítis**	**appéndix**	**#LHL**
77. apprèciátory[13]	appréciate	#LHL
78. appròachabílity	appróachable	#LHL
79. appròximátion	appróximate	#LHL
80. assòciátion	assóciate	#LHL
81. attàinabílity	attáinable	#LHL
82. avàilabílity	aváilable	#LHL
83. **Banànaráma**	**banána**	**#LHL**
84. **caprìccióso**	**capríccio**	**#LHL**
85. circùmferéntial	circúmference	#LHL
86. collàpsibílity	collápsible	#LHL
87. collèctivístic	colléctivist	#LHL

[12] Also *abbréviatory*.
[13] Also *appréciatory*.

88. collègiálity	collégial	#LHL
89. commùnicátion	commúnicate	#LHL
90. commùtabílity	commútable	#LHL
91. corrùptibílity	corrúptible	#LHL
92. kalèidoscópic	kaléidoscope	#LHL
93. patèrnalístic	patérnal	#LHL
94. percèptibílity	percéptible	#LHL
95. **pronùnciaménto**	**pronóunce**	**#LHL**
96. pronùnciátion	pronóunce	#LHL
97. propòrtionálity	propórtional	#LHL
98. provìnciálity	províncial	#LHL
99. substàntiátion	substántiate	#LHL
100. suscèptibílity	suscéptible	#LHL
101. sustàinabílity	sustáinability	#LHL

Appendix H: #LHL and #LHH words from Jones (2003)

#LHL and #LHH words

Embedding word	Embedded word	
1. accùlturátion	accúlturate	#LHL
2. **aràchnophóbia**	**aráchnid (OED)**	**#LHL/H**
3. deàctivátion	deáctivate	#LHL
4. dedùcibílity	dedúcible	#LHL
5. defènsibílity	defénsible	#LHL
6. denùnciátion	denúnciate	#LHL
7. depèndabílity	depéndable	#LHL
8. deprèciátion	depréciate	#LHL
9. desìrabílity	desírable	#LHL
10. despòliátion	despóil	#LHL
11. destrùctibílity	destrúctible	#LHL
12. detèrminátion	detérminate	#LHL
13. diffùsibílity	diffúsible	#LHL
14. dilàtabílity	dilátable (OED)	#LHL
15. dissòciátion	dissóciate	#LHL
16. dissòlvabílity	dissólve	#LHL
17. **divèrtiménto**	**divért**	**#LHL**
18. **Ecclèsiástes**	**Ecclésiast (OED)**	**#LHL**
19. ecclèsiástic	Ecclésiast (OED)	#LHL
20. effèctuálity	efféctual	#LHL
21. elùcidátion	elúcidate	#LHL
22. emàncipátion	emáncipate	#LHL
23. enùmerátion	enúmerate	#LHL
24. enùnciátion	enúnciate	#LHL
25. essèntiálity	esséntial	#LHL
26. evèntuálity	evéntual	#LHL
27. **illùmináti**	**illúminate**	**#LHL**
28. illùminátion	illúminate	#LHL
29. immòvabílity	immóvable	#LHL
30. inàugurátion	ináugurate	#LHL
31. inèbriátion	inébriate	#LHL
32. ***listeriosis***	***listéria***	***#LHL***
33. negòtiátion	negótiate	#LHL
34. ostènsibílity	osténsible	#LHL
35. pecùliárity	pecúliar	#LHL
36. prevèntabílity	prevéntable (OED)	#LHL
37. **pronùnciaménto**	**pronóunce**	**#LHL**
38. recùperátion	recúperate	#LHL
39. redìntegrátion	redíntegrate (OED)	#LHL
40. redùcibílity	redúcible	#LHL
41. redùplicátion	redúplicate	#LHL
42. regùrgitátion	regúrgitate	#LHL
43. rejùvenátion	rejúvenate	#LHL

44. remèdiátion	remédiate (OED)	#LHL
45. remòvabílity	remóvable (OED)	#LHL
46. remùnerátion	remúnerate	#LHL
47. renùnciátion	renóunce	#LHL
48. repèatabílity	repéatable (OED)	#LHL
49. repùdiátion	repúdiate	#LHL
50. resòlvabílity	resólvable (OED)	#LHL
51. respèctabílity	respéctable	#LHL
52. respònsibílity	respónsible	#LHL
53. revèrberátion	revérberate	#LHL
54. revèrsibílity	revérsible	#LHL
55. torrèntiálity	torréntial	#LHL

#L(=R)HL and #LHH words

56. abbrèviátion	abbréviate	#LHL
57. accòuntabílity	accóuntable	#LHL
58. accùmulátion	accúmulate	#LHL
59. accùsatíval	accúsative	#LHL
60. accùsatívity	accúsative	#LHL
61. adàptabílity	adáptable	#LHL
62. adjùdicátion	adjúdicate	#LHL
63. adùlterátion	adúlterate	#LHL
64. agglùtinátion	agglútinate	#LHL
65. allèviátion	alléviate	#LHL
66. amàlgamátion	amálgamate	#LHL
67. amènabílity	aménable	#LHL
68. annìhilátion	anníhilate	#LHL
69. annùnciátion	annúnciate	#LHL
70. **appèndicítis**	**appéndix**	**#LHL**
71. apprèciátion	appréciate	#LHL
72. appròachabílity	appróachable	#LHL
73. appròpriátion	appróориate	#LHL
74. appròximátion	appróximate	#LHL
75. assòciátion	assóciate	#LHL
76. attàinabílity	attáinable (OED)	#LHL
77. attràctabílity	attráctable (OED)	#LHL
78. avàilabílity	aváilable	#LHL
79. calùmniátion	calúmniate	#LHL
80. commèrciálity	commércial	#LHL
81. commùnicátion	commúnicate	#LHL
82. commùtabílity	commútable	#LHL
83. corrùptibílity	corrúptible	#LHL
84. hallùcinátion	hallúcinate	#LHL
85. kalèidoscópic	kaléidoscope	#LHL
86. Napòleónic	Napóleon	#LHL
87. peràmbulátion	perámbulate	#LHL
88. percèptibílity	perceptible	#LHL
89. predìctabílity	predíctable	#LHL

90. pronùnciátion	pronóunce	#LHL
91. propòrtionálity	propórtional	#LHL
92. proprìetórial	propríetor	#LHL
93. subòrdinátion	subórdinate	#LHL
94. suscèptibílity	suscéptible	#LHL
95. suspènsibílity	suspénse	#LHL
96. sustàinabílity	sustáinable (OED)	#LHL

Appendix I: #LLL and #LLH words from Wells (2000) where initial-syllable preservation is expected

#LLL words

Embedding word	Embedded word		
1. àmicabílity	ámicable		#LLL
2. ànimalístic	ánimal		#LLL
3. Àristotélian	Áristotle	Rem.	#LLL
4. cànnibalístic	cánnibal	Rem.	#LLL
5. càpitalístic	cápitalist		#LLL
6. clàrificátion	clárify		#LLL
7. clàssificátion	clássify	Rem.	#LLL
8. clàssificátory[14]	clássify	Rem.	#LLL
9. dìsciplinárian	díscipline		#LLL
10. èdificátion	édify		#LLL
11. èducabílity	éducable		#LLL
12. èligibílity	éligible		#LLL
13. èpigrammátic	épigram	Rem.	#LLL
14. gàsificátion	gásify		#LLL
15. glàdiatórial	gládiator		#LLL
16. gràtificátion	grátify		#LLL
17. hàbitabílity	hábitable		#LLL
18. ìrritabílity	írritable	Rem.	#LLL
19. jòllificátion	jóllify (OED)	Rem.	#LLL
20. jùstificátion	jústify	Rem.	#LLL
21. màlleabílity	málleable	Rem.	#LLL
22. mànageabílity	mánageable		#LLL
23. màrriageabílity	márriageable	Rem.	#LLL
24. mèlodramátic	mélodràma	Rem.	#L.L/H.L
25. mòdificátion	módify		#LLL
26. mòllificátion	móllify	Rem.	#LLL
27. mùmmificátion	múmmify	Rem.	#LLL
28. mỳstificátion	mýstify	Rem.	#LLL
29. nàvigabílity	návigable		#LLL
30. nùllificátion	núllify	Rem.	#LLL
31. òssificátion	óssify	Rem.	#LLL
32. pàlatabílity	pálatable		#LLL
33. pènetrabílity	pénetrable		#LLL
34. pèregrinátion	péregrinate		#LLL
35. pèrishabílity	périshable		#LLL
36. prèdicabílity	prédicable		#LLL
37. prèttificátion	préttify	Rem.	#LLL
38. pròfitabílity	prófitable		#LLL
39. qùalificátion	qúalify		#LLL
40. qùalificátive	qùalify		#LLL

[14] Also *clássificatory*.

41. ràmificátion	rámify		#LLL
42. ràtificátion	rátify		#LLL
43. rèplicabílity	réplicable		#LLL
44. rèputabílity	réputable		#LLL
45. rìtualístic	rítualist		#LLL
46. Rùssificátion	Rússify	Rem.	#LLL
47. spècificátion	spécify		#LLL
48. spìritualístic	spíritualist		#LLL
49. spìrituálity	spíritual		#LLL
50. stràtificátion	strátify		#LLL
51. ùglificátion	úglify		#LLL
52. vèrificátion	vérify		#LLL
53. vìlificátion	vílify		#LLL
54. vìvificátion	vívify		#LLL
55. **Wìnnipegósis**	**Wínnipeg**	**Rem.**	**#LLL**
56. yùppificátion	yúppify	Rem.	#LLL

#LL(=R)L words

57. ìteratívity	íterative
58. nàtionalístic	nátionalist
59. nàturalístic	náturalist
60. nègativístic	négative
61. òperabílity	óperable
62. Pèloponnésian	Péloponnese[15]
63. pròbabilístic	próbable
64. ràtionalístic	rátionalist
65. rèlativístic	rélativist (OED)
66. sèparabílity	séparable
67. **Trìpolitánia**	**Trípoli**

Sometimes or always #LLH words

68. àlimentátion	áliment (OED)
69. cànalizátion	cánalize
70. cànonizátion	cánonize
71. cìvilizátion	cívilize
72. clìticizátion	clíticize
73. còlonizátion	cólonize
74. crỳstallizátion	crýstallize
75. dìgitizátion	dígitize
76. dòcumentátion	dócument
77. **èlephantíasis**	**élephant**
78. fèminizátion	féminize
79. fòssilizátion	fóssilize
80. fràternizátion	fráternize
81. glàmorizátion	glámorize
82. glòttalizátion	glóttalize

[15] Also *Pèloponnése*.

83.	ìmmunizátion	ímmunize
84.	**làbyrinthítis**	**lábyrinth**
85.	làtinizátion	látinize
86.	lèmmatizátion	lémmatize
87.	mèritocrátic	méritocrat
88.	mèmorizátion	mémorize
89.	mòdernizátion	módernize
90.	pàlatográphic	pálatograph (OED)
91.	pàlletizátion	pálletize
92.	**pàrasitólogy**	**párasite**
93.	**pàronomásia**	**páronym**
94.	pàssivizátion	pássivize
95.	phànerogámic	pháneragam
96.	pìdginizátion	pídginize
97.	plàsticizátion	plásticize
98.	ràndomizátion	rándomize
99.	règimentátion	régiment
100.	Rùssianizátion	Rússianize
101.	sàtirizátion	sátirize
102.	sèdimentátion	sédiment
103.	spècializátion	spécialize
104.	stèrilizátion	stérilize
105.	sùpplementátion	súpplement
106.	vàlorizátion	válorize
107.	wèsternizátion	wésternize

Appendix J: #LLL and #LLH words from Jones (2003) where initial-syllable preservation is expected

#LLL words

Embedding word	Embedded word		
1. àmicabílity	ámicable		#LLL
2. ànimalístic	ánimal		#LLL
3. Àristotélian	Áristotle	Rem.	#LLL
4. cànnibalístic	cánnibal	Rem.	#LLL
5. clàrificátion	clárify		#LLL
6. clàssificátion	clássify	Rem.	#LLL
7. clàssificátory	clássify	Rem.	#LLL
8. dìsciplinárian	díscipline		#LLL
9. èdificátion	édify		#LLL
10. èducabílity	éducable		#LLL
11. èligibílity	éligible		#LLL
12. èpigrammátic	épigram	Rem.	#LLL
13. fìgurabílity	fígurable		#LLL
14. glàdiatórial	gládiator		#LLL
15. gràtificátion	grátify		#LLL
16. ìmitabílity	ímitable		#LLL
17. ìrritabílity	írritable	Rem.	#LLL
18. jòllificátion	jóllify	Rem.	#LLL
19. jùstificátion	jústify	Rem.	#LLL
20. jùstificátory	jústify	Rem.	#LLL
21. màlleabílity	málleable	Rem.	#LLL
22. mànageabílity	mánageable		#LLL
23. mìlitarístic	mílitary		#LLL
24. mòdificátion	módify		#LLL
25. mòllificátion	móllify	Rem.	#LLL
26. mùmmificátion	múmmify	Rem.	#LLL
27. mỳstificátion	mýstify	Rem.	#LLL
28. nàturalistic	náturalist (OED)		#LLL
29. nàvigabílity	návigable		#LLL
30. nègligibílity	négligible		#LLL
31. nùllificátion	núllify	Rem.	#LLL
32. òssificátion	óssify	Rem.	#LLL
33. pàcificátion	pácify		#LLL
34. pènetrabílity	pénetrable		#LLL
35. pèregrinátion	péregrinate		#LLL
36. pèrishabílity	périshable		#LLL
37. prèdicabílity	prédicable		#LLL
38. prèttificátion	préttify	Rem.	#LLL
39. pròfitabílity	prófitable		#LLL
40. qùalificátion	qúalify		#LLL
41. qùalificátory	qúalify		#LLL
42. qùalificátive	qúalify		#LLL

43. ràmificátion	rámify	#LLL
44. ràtificátion	rátify	#LLL
45. rèputabílity	réputable	#LLL
46. rìtualístic	rítual	#LLL
47. scàrificátion	scárify	#LLL
48. spècificátion	spécify	#LLL
49. spìrituálity	spíritual	#LLL
50. ***Tripolitania***	***Trípoli***	***#LLL***
51. ùglificátion	úglify	#LLL
52. vàriegátion	váriegate	#LLL
53. vèrificátion	vérify	#LLL
54. vìlificátion	vílify	#LLL
55. vìtrificátion	vítrify	#LLL
56. vìvificátion	vívify	#LLL

#LL(=R)L words

57. nàtionalístic	nátionalist (OED)
58. òperabílity	óperable
59. pàronomásia	**páronym**
60. Pèloponnésian	Péloponnese[16]
61. pròbabilístic	próbable
62. ràtionalitic	rátional
63. rèlativistic	rélativist (OED)

Sometimes or always #LLH words

64. cànalizátion	cánalize
65. cànonizátion	cánonize
66. chànnelizátion	chánnelize
67. cìvilizátion	cívilize
68. còlonizátion	cólonize
69. crỳstallizátion	crýstallize
70. dìgitizátion	dígitize
71. dòcumentátion	dócument
72. èlephantíasis	**élephant**
73. fòssilizátion	fóssilize
74. fràternizátion	fráternize
75. ìdeográphic	ídeograph
76. ìdiomátic	ídiom
77. ìmmunizátion	ímmunize
78. mànifestátion	mánifest
79. mèchanizátion	méchanize
80. mìnimizátion	mínimize
81. mòdernizátion	módernize
82. nòvelizátion	nóvelize
83. **pàrasitólogy**	**párasite**
84. pòlarizátion	pólarize

[16] Also *Pèloponnése*.

85.	rèalizátion	réalize
86.	règimentátion	régiment
87.	sòlemnizátion	sólemnize
88.	spècializátion	spécialize
89.	stèrilizátion	stérilize
90.	tèrrorizátion	térrorize
91.	vàlorizátion	válorize
92.	wèsternizátion	wésternize

Appendix K: BNC frequency counts for -*ation* words from Jones (2003) where relative prominence preservation is expected

-*ation* words: preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
1. accèlerátion	7	accélerate	11
2. accèntuátion	0	accéntuate	4
3. artìculátion	4	artículate	8
4. asphỳxiátion	0	asphýxiate	0
5. authènticátion	0	authénticate	1
6. coàgulátion	1	coágulate	0
7. concìliátion	2	concíliate	0
8. confàbulátion	0	confábulate	0
9. confèderátion	6	conféderate	0
10. congràtulátion	7	congrátulate	9
11. consideration	78	consíder	289
12. consòlidátion	5	consólidate	9
13. contàminátion	6	contáminate	5
14. continuation	9	contínue	283
15. coòrdinátion	5	coórdinate	6
16. defìbrillátion	0	defíbrillate	0
17. ejàculátion	1	ejáculate	0
18. elìcitátion	0	elícit	6
19. emàciátion	0	emáciate	0
20. emàsculation	0	emásculate	0
21. encàpsulátion	0	encápsulate	3
22. evìscerátion	0	evíscerate	0
23. exàcerbátion	0	exácerbate	6
24. exàggerátion	4	exággerate	9
25. exàminátion	62	exámine	95
26. exàsperátion	3	exásperate	2
27. excòriátion	0	excóriate	0
28. excrùciátion	0	excrúciate	*0*
29. exhìlarátion	1	exhílarate	0
30. exònerátion	0	exónerate	1
31. expàtiátion	0	expátiate	0
32. expèctorátion	0	expéctorate	0
33. expòstulátion	0	expóstulate	0
34. extènuátion	0	exténuate	0
35. extèrminátion	1	extérminate	1

36. extràpolátion	1	extrápolate	2
37. exùviátion	0	exúviate	*0*
38. hypòthecátion	0	hypóthecate	0
39. impèrsonátion	1	impérsonate	1
40. incìnerátion	2	incínerate	1
41. incòrporátion	5	incórporate	43
42. incrìminátion	0	incríminate	1
43. indòctrinátion	1	indóctrinate	0
44. infàtuátion	1	infátuate	0
45. infìbulátion	0	infíbulate	0
46. inhàbitátion	0	inhábit	7
47. insèminátion	1	inséminate	0
48. insìnuátion	1	insínuate	1
49. intèrcalátion	0	intércalate	0
50. intèrpolátion	1	intérpolate	0
51. intèrpretátion	54	intérpret	43
52. intèrrogátion	4	intérrogate	3
53. intìmidátion	3	intímidate	4
54. intòxicátion	1	intóxicate	1
55. invàlidátion	0	inválidate	2
56. invèstigátion	68	invéstigate	55
57. invìgilátion	0	invígilate	0
58. invìgorátion	0	invígorate	0
59. perpètuátion	1	perpétuate	4
60. persèverátion	0	persévérate	0
61. prepònderátion	0	prepónderate	0
62. procràstinátion	1	procrástinate	0
63. prognòsticátion	0	prognósticate	0
64. prolìferátion	7	prolíferate	2
65. reòccupátion	0	reóccupy	0
66. substàntiátion	0	substántiate	3
67. vitùperátion	0	vitúperate	0
68. vocìferátion	0	vocíferate	0

-ation words: non-preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
69. anticipation	8	antícipate	24
70. cohabitation	1	cohábit	1
71. concatenation	0	concátenate	0
72. concelebration	0	concélebrate	0

73. configuration	13	confígure	2
74. conglomeration	1	conglómerate	0
75. cooperation	12	coóperate	5
76. de-escalation	0	deéscalate	0
77. defoliation	0	defóliate	0
78. deforestation	2	defórest	0
79. depopulátion	1	depópulate	0
80. deracination	0	derácinate	0
81. desalination	0	desálinate	0
82. desegragation	0	deségregate	0
83. exfoliation	0	exfóliate	0
84. expropriation	1	exprópriate	1
85. humiliation	6	humíliate	4
86. impropriation	0	imprópriate	0
87. incarcerátion	1	incárcerate	1
88. participation	27	partícipate	29
89. prefabrication	0	prefábricate	0
90. premeditation	0	preméditate	0
91. preoccupation	8	preóccupy	1
92. transliteration	0	translíterate	0
93. dèconsecrátion	0	decónsecrate	0
94. dèregulátion	4	derégulate	1
95. prèdestinátion	1	predéstine	0
96. rèactivation	0	reáctivate	1
97. règenerátion	6	regénerate	2
98. rèpatriátion	4	repátriate	1

Appendix L: CELEX frequency counts for -*ation* words from Jones (2003) where relative prominence preservation is expected

-*ation* words: preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
1. accèlerátion	143	accélerate	269
2. accèntuátion	3	accéntuate	71
3. artìculátion	28	artículate	63
4. asphỳxiátion	2	asphýxiate	2
5. authènticátion	2	authénticate	16
6. coàgulátion	3	coágulate	11
7. concìliátion	30	concíliate	5
8. confàbulátion	2	confábulate	0
9. confèderátion	64	conféderate	5
10. congràtulátion	36	congrátulate	163
11. consideration	716	consíder	3730
12. consòlidátion	45	consólidate	128
13. contàminátion	80	contáminate	96
14. continuation	119	contínue	4702
15. coòrdinátion	110	coórdinate	261
16. defìbrillátion	0	defíbrillate	0
17. ejàculátion	15	ejáculate	7
18. elìcitátion	0	elícit	75
19. emàciátion	2	emáciate	25
20. emàsculation	2	emásculate	9
21. encàpsulátion	0	encápsulate	31
22. evìscerátion	0	evíscerate	5
23. exàcerbátion	1	exácerbate	69
24. exàggerátion	104	exággerate	231
25. exàminátion	1259	exámine	1387
26. exàsperátion	69	exásperate	79
27. excòriátion	1	excóriate	1
28. excrùciátion	0	excrúciate	0
29. exhìlarátion	41	exhílarate	89
30. exònerátion	3	exónerate	23
31. expàtiátion	0	expátiate	1
32. expèctorátion	0	expéctorate	0
33. expòstulátion	6	expóstulate	17
34. extènuátion	1	exténuate	4
35. extèrminátion	31	extérminate	41

36. extràpolátion	0	extrápolate	17
37. exùviátion	0	exúviate	0
38. hypòthecátion	0	hypóthecate	0
39. impèrsonátion	24	impérsonate	22
40. incìnerátion	16	incínerate	14
41. incòrporátion	25	incórporate	295
42. incrìminátion	13	incríminate	48
43. indòctrinátion	20	indóctrinate	18
44. infàtuátion	27	infátuate	0
45. infìbulátion	0	infíbulate	0
46. inhàbitátion	0	inhábit	206
47. insèminátion	5	inséminate	5
48. insìnuátion	12	insínuate	32
49. intèrcalátion	0	intércalate	0
50. intèrpolátion	6	intérpolate	6
51. intèrpretátion	463	intérpret	459
52. intèrrogátion	97	intérrogate	73
53. intìmidátion	48	intímidate	96
54. intòxicátion	13	intóxicate	42
55. invàlidátion	1	inválidate	24
56. invèstigátion	511	invéstigate	412
57. invìgilátion	0	invígilate	1
58. invìgorátion	0	invígorate	32
59. perpètuátion	19	perpétuate	101
60. persèverátion	0	persévérate	0
61. prepònderátion	0	prepónderate	1
62. procràstinátion	5	procrástinate	6
63. prognòsticátion	4	prognósticate	0
64. prolìferátion	90	prolíferate	49
65. reòccupátion	0	reóccupy	0
66. substàntiátion	1	substántiate	27
67. vitùperátion	1	vitúperate	0
68. vocìferátion	2	vocíferate	0

-ation words: non-preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
69. anticipation	133	antícipate	382
70. cohabitation	17	cohábit	7
71. concatenation	6	concátenate	0
72. concelebration	0	concélebrate	0

73. configuration	41	confígure	0
74. conglomeration	14	conglómerate	0
75. cooperation	446	coóperate	202
76. de-escalation	12	deéscalate	6
77. defoliation	4	defóliate	3
78. deforestation	93	defórest	5
79. depopulátion	7	depópulate	10
80. deracination	0	derácinate	0
81. desalination	2	desálinate	1
82. desegregation	1	deségregate	1
83. exfoliation	0	exfóliate	0
84. expropriation	15	exprópriate	17
85. humiliation	188	humíliate	199
86. impropriation	0	imprópriate	0
87. incarceration	20	incárcerate	16
88. participation	352	partícipate	386
89. prefabrication	1	prefábricate	0
90. premeditation	10	preméditate	5
91. preoccupation	193	preóccupy	23
92. transliteration	4	translíterate	0
93. dèconsecrátion	0	decónsecrate	0
94. dèregulátion	0	derégulate	0
95. prèdestinátion	10	predéstine	10
96. rèactivation	0	reáctivate	10
97. règenerátion	61	regénerate	47
98. rèpatriátion	40	repátriate	14

Appendix M: CELEX frequency counts for -*ity* words from Jones (2003) where relative prominence preservation is expected

-*ity* words: preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
1. acceptability	17	acceptable	518
2. accessibility	27	accessible	164
3. admissibility	0	admissible	7
4. advisability	10	advisable	76
5. bisexuality	1	bisexual	7
6. combustibility	0	combustible	14
7. compatibility	17	compatible	102
8. compressibility	1	compressible	2
9. conductibility	0	conductible	0
10. conformability	0	conformable	1
11. congeniality	0	congenial	45
12. connubiality	0	connubial	7
13. contemptibility	0	contemptible	24
14. contractibility	0	contractible	0
15. conventionality	5	conventional	834
16. convertibility	4	convertible	27
17. conviviality	8	convivial	26
18. digestibility	4	digestible	12
19. dilatability	0	dilatable	0
20. exchangeability	0	exchangeable	4
21. excitability	2	excitable	24
22. expansibility	0	expansible	0
23. extensibility	0	extensible	0
24. impersonality	0	impersonal	121
25. implacability	0	implacable	55
26. impregnability	1	impregnable	43
27. impressibility	0	impressible	0
28. improvability	0	improvable	0
29. incapability	1	incapable	259
30. incredibility	1	incredible	347
31. incurability	0	incurable	34
32. indelibility	0	indelible	20
33. infallibility	7	infallible	40
34. infrangibility	0	infrangible	0
35. insatiability	0	insatiable	55
36. inscrutability	3	inscrutable	32

37. insensibility	4	insensible	15
38. insensitivity	18	insensitive	63
39. insolubility	0	insoluble	55
40. intangibility	1	intangible	23
41. intractability	3	intractable	45
42. intransitivity	0	intransitive	0
43. invincibility	10	invincible	10
44. invisibility	10	invisible	425
45. perfectibility	1	perfectible	1
46. potentiality	66	potential	613
47. subsidiarity	0	subsidiary	161
48. substantiality	0	substantial	599
49. unchangeability	0	unchangeable	0
50. unpopularity	0	unpopular	0
51. unpracticality	0	unpractical	0
52. unpunctuality	0	unpunctual	0
53. untouchability	0	untouchable	20

-ity words: non-preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
54. foreseeability	0	foreseeable	55
55. impartiality	32	impartial	47
56. impassability	0	impassability	22
57. impeccability	0	impeccable	56
58. impossibility	98	impossible	1865
59. implausibility	0	implausible	0
60. impracticality	2	impractical	47
61. improbability	10	improbable	123
62. imputability	0	imputable	0
63. municipality	8	municipal	120
64. transmissibility	0	transmissible	0
65. transmutability	0	transmutable	2
66. unsuitability	0	unsuitable	0

Appendix N: CELEX frequency counts for #LLL -*ion* words from Jones (2003) where second-syllable preservation is expected

#LLL -*ion* words: preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
1. abòminátion	24	abóminate	5
2. accòmmodátion	607	accómmodate	263
3. acètylátion	0	acétylate	0
4. agglòmerátion	7	agglómerate	0
5. allìterátion	0	allíterate	0
6. assàssinátion	137	assássinate	64
7. assèverátion	1	asséverate	0
8. assìbilátion	0	assíbilate	0
9. assìmilátion	31	assímilate	113
10. attènuátion	3	atténuate	21
11. capìtulátion	18	capítulate	32
12. collàborátion	102	colláborate	72
13. commèmorátion	9	commémorate	57
14. commìserátion	10	commíserate	0
15. deàminátion	0	deáminate	0
16. debìlitátion	0	debílitate	41
17. decàpitátion	5	decápitate	22
18. degènerátion	18	degénerate	53
19. delìberátion	104	delíberate	40
20. delìneátion	8	delíneate	25
21. demòdulátion	0	demódulate	0
22. denòminátion	36	denóminate	1
23. desìderátion	0	desíderate	0
24. dilàpidátion	5	dilápidate	0
25. discrìminátion	287	discríminate	115
26. dissèminátion	41	disséminate	48
27. dissìmilátion	0	dissímilate	0
28. dissìmulátion	7	dissímulate	0
29. domèsticátion	31	doméstícate	70
30. elàborátion	54	eláborate	90
31. elìminátion	106	elíminate	539
32. emàsculátion	2	emásculate	9
33. eqùivocátion	12	equívocate	12
34. eràdicátion	21	erádicate	56
35. evàcuátion	49	evácuate	60
36. evàluátion	169	eváluate	151

37. evìscerátion	0	evíscerate	5
38. facìlitátion	5	facílitate	126
39. felìcitátion	0	felícitate	0
40. gestìculátion	20	gestículate	46
41. habìlitátion	0	habílitate	0
42. immòderátion	0	immóderate	6
43. inòculátion	41	inóculate	17
44. inòsculátion	0	inósculate	0
45. irrèsolútion	2	irrésolute	9
46. legìtimátion	0	legítimate	0
47. manìpulátion	128	manípulate	237
48. matrìculátion	28	matrículate	8
49. oblìterátion	13	oblíterate	83
50. orìginátion	0	oríginate	175
51. precìpitátion	12	precípitate	75
52. predèstinátion	10	predéstinate	0
53. predòminátion	0	predóminate	60
54. premèditátion	10	preméditate	5
55. prevàricátion	4	prevâricate	10
56. propìtiátion	9	propítiate	9
57. reàllocátion	0	reállocate	0
58. recìprocátion	1	recíprocate	34
59. refrìgerátion	15	refrígerate	21
60. regènerátion	61	regénerate	47
61. repàtriátion	40	repátriate	14
62. resùscitátion	5	resúscitate	10
63. retàliátion	100	retáliate	92
64. retìculátion	1	retículate	0
65. sophìsticátion	105	sophísticate	0
66. syllàbicátion	0	syllábicate	0

#LLL *-ion* words: non-preserving stress			
Embedding word	Embedding frequency	Embedded word	Embedded frequency
67. horripilation	0	horrípilate	0
68. mìscegenation (init.)	6	miscégenate	0
69. vaticination	0	vatícinate	0

References

Alber, B. (1997). “Quantity sensitivity as the result of constraint interaction”. In G. Booij & J. van der Weijer (eds.), *Phonology in progress – progress in phonology* [HIL phonology papers III]. The Hague: Holland Academic Press. 1-45.

Alber, B. (1998). “Stress preservation in German loan-words”. In W. Kehrein & R. Wiese (eds.), *Phonology and morphology of the Germanic languages*. Tübingen: Niemeyer. 113-141.

Anshen, F. & M. Aronoff (1988). “Producing morphologically complex words”. *Linguistics* 26: 641-655.

Anttila, A. (1997). “Variation in Finnish phonology and morphology”. PhD dissertation, Stanford University, Stanford, CA.

Anttila, A. (2002a). “Morphologically conditioned phonological alternations”. *Natural Language and Linguistic Theory* 20: 1-41.

Anttila, A. (2002b). “Variation and phonological theory”. In J. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *Handbook of language variation and change*. Oxford: Blackwell. 206-243.

Anttila, A. (2007). “Variation and optionality”. In Paul de Lacy, (ed.), *The Cambridge handbook of phonology*. Cambridge: CUP. 519-536.

Anttila, A. & Y.Y. Cho (1998). “Variation and change in Optimality Theory”. *Lingua* 104: 31-56.

Arbib, M.A. (1987). “Levels of modelling of mechanisms of visually guided behavior”. *The Behavioral and Brain Sciences* 10: 407-465.

Archangeli, D. (1984). “An overview of the theory of Lexical Phonology and Morphology”. *MIT Working Papers in Linguistics* 7: 1-14.

Aronoff, M. (1976). *Word formation in Generative Grammar*. Cambridge, MA: MIT Press.

Aronoff, M. & S.N. Sridhar (1983). “Morphological levels in English and Kannada, or Atarizing Reagan”. *Papers from the parasession on the interplay of phonology, morphology and syntax*. Chicago: Chicago Linguistics Society. 3-16.

Aronoff, M. & F. Anshen (1998). “Morphology and the lexicon: lexicalization and productivity”. In A. Spencer & A.M. Zwicky (eds.), *The handbook of morphology*. Oxford: Blackwell. 237-247.

Baayen, H. (1992). “Quantitative aspects of morphological productivity”. In G. Booij & J. van Marle (eds.), *Yearbook of morphology 1991*. Dordrecht: Foris. 109-149.

Baayen, R.H., R. Piepenbrock & L. Gulikers (1995). *The CELEX Lexical Database* (Release 2) [CD-ROM]. English version 2.5. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Baayen, R.H. & R. Schreuder (2000). “Towards a psycholinguistic computational model for morphological parsing”. *Philosophical Transactions of the Royal Society (Series A: Mathematical, Physical and Engineering Sciences)* 358: 1-13.

Baayen, R.H., R. Schreuder & R. Sproat (2000). “Morphology in the mental lexicon: a computational model for visual word recognition”. In F. van Eynde & D. Gibbon (eds.), *Lexicon development for speech and language processing*. Dordrecht: Kluwer Academic Publishers. 267-291.

Badecker, W. (1991). “Affix raising and the level ordering hypothesis”. *Lingua* 83: 103-32.

Baković, E.J. (1995). “Strong onsets in Spanish fortition”. Ms., Rutgers University. [ROA-96-0096, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Balogné Bérces, K. (2006). “What’s wrong with vowel-initial syllables?” *SOAS Working Papers in Linguistics* 14: 15-21.

Bauer, L. (1988). *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press.

Benua, L. (1995). “Identity effects in morphological truncation”. Ms., University of Massachusetts, Amherst. Draft of July 1995. [ROA-74-0000, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Benua, L. (1997). *Transderivational identity: phonological relations between words*. PhD dissertation, University of Massachusetts, Amherst. [ROA-259-0498, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Bermúdez-Otero, R. (1999). *Constraint interaction in language change: quantity in English and Germanic [Opacity and globality in phonological change]*. PhD dissertation, University of Manchester & Universidad de Santiago de Compostela. [URL: www.bermudez-otero.com/PhD.pdf (11:55, 13th June 2007)]

Bermudez-Otero, R. (2002). “Cyclicity or sympathy? A case study.” Handout of paper presented at the *10th Manchester Phonology Meeting*, University of Manchester, 25th May 2002. [URL: www.bermudez-otero.com/10mfm.pdf (11:50, 13th June 2007)]

Bermúdez-Otero, R. (2003). "The acquisition of phonological opacity". Ms., University of Manchester.
[ROA-593-0403, Rutgers Optimality Archive, http://roa.rutgers.edu/]
[Shorter version published in J. Spenader, A. Eriksson & Ö. Dahl (eds), *Variation within Optimality Theory: proceedings of the Stockholm workshop on variation within Optimality Theory´*. Stockholm: Department of Linguistics, Stockholm University. 25-36.]

Bermúdez-Otero, R. (2007a). "Morphological structure and phonological domains in Spanish denominal derivation". In F. Martínez-Gil & S. Colina (eds.), Optimality-theoretic studies in Spanish phonology. Amsterdam: John Benjamins. 278–311.

Bermúdez-Otero, R. (2007b). "On the nature of the cycle". Handout of paper presented at the special session "Where is allomorphy?", *15th Manchester Phonology Meeting*, University of Manchester, 25th May 2007.
[URL: www.bermudez-otero.com/15mfm.pdf (11:53, 13th June 2007)]

Bermúdez-Otero, R. (in preparation). *Stratal Optimality Theory*. Cambridge: OUP.

Bermúdez-Otero, R. & A.M.S. McMahon (2006). "English phonology and morphology". In B. Aarts & A.M.S. McMahon (eds.), *The handbook of English linguistics*. Oxford: Blackwell. 382-410.

Black, H.A. (1994). *Constraint-ranked derivation: a serial approach to optimization*. PhD dissertation, University of California, Santa Cruz.

Bolinger, D.L. (1961). "Contrastive accent and contrastive stress". *Language* 37/1: 83-96.

Booij, G.E. (1977). *Dutch morphology: a study of word formation in generative grammar*. Lisse: Pieter de Ridder Press.

Booij, G.E. (1996). "Lexical phonology and the derivational residue". In J. Durand & B. Laks (eds.), *Current trends in phonology: models and methods*. Volume 1. Salford: European Studies Research Institute. 69-96.

Booij, G.E. (1997). "Non-derivational phonology meets Lexical Phonology". In I. Roca (ed.), *Derivations and constraints in phonology*. Oxford: Clarendon Press. 261-288.

Booij, G.E. & J. Rubach (1984). "Morphological and prosodic domains in Lexical Phonology". *Phonology Yearbook* 1: 1-27.

Booij, G.E. & J. Rubach (1987). "Postcyclic vs. post-lexical rules in Lexical Phonology". *Linguistic Inquiry* 18: 1-44.

Borowsky, T. (1993). "On the word level". In S. Hargus & E.M. Kaisse (eds.), *Phonetics and phonology: studies in Lexical Phonology*. Volume 4. London: Academic Press, Inc.

Brame, M. (1974). "The cycle in phonology: stress in Palestinian, Maltese and Spanish". *Linguistic Inquiry* 5: 39-60.

Brown, A.S. (1991). "A review of the tip-of-the-tongue experience". *Psychological Bulletin* 109/2: 204-223.

Burzio, L. (1994). *Principles of English stress*. Cambridge: CUP.

Burzio, L. (1996). "Surface constraints versus underlying representations". In J. Durand & B. Laks (eds.), *Current trends in phonology: models and methods*. Volume 1. Salford: European Studies Research Institute. 123-142.

Burzio, L. (1999). "Surface-to-surface morphology: when your representations turn into constraints". Ms., John Hopkins University. [ROA-341-0999, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Burzio, L. (2000). "Cycles, non-derived-environment blocking, and correspondence". In J. Dekkers, F, van der Leeuw, & J. van de Weijer (eds.), *Optimality Theory: phonology, syntax and acquisition*. Oxford: OUP. 47-87.

Burzio, L. (2002). "Phonology and phonetics of English stress and vowel reduction". Ms., Department of Cognitive Science, John Hopkins University, Baltimore.

Burzio, L. (2005a). "Lexicon and grammar: unequal but inseparable". Ms., Department of Cognitive Science, John Hopkins University, Baltimore.

Burzio, L. (2005b). "Sources of paradigm uniformity". In L.J. Downing, T.A. Hall & R. Raffelsiefen (eds.), *Paradigms in phonological theory*. Oxford: OUP. 65-106.

Burzio, L. (2007). "Phonology and phonetics of English stress and vowel reduction". *Language Sciences* 29: 154-176. [URL: http://web.jhu.edu/cogsci/templates/images/pdf/burzio/Phonology%20and%20Phonetics.pdf (15:44, 13th June 2007)]

Butterworth, B. (1983). "Lexical representation". In B. Butterworth (ed.), *Language Production*. Volume 2. London: Academic Press. 257-294.

Bybee, J. (1985). *Morphology: a study of the relations between meaning and form*. Amsterdam: Benjamins.

Caramazza, A., A. Laudanna & C. Romani (1988). "Lexical access and inflectional morphology". *Cognition* 28: 297-332.

Chomsky, N. (1970). “Remarks on nominalization”. In R.A. Jacobs & P.S. Rosenbaum (eds.), *Readings in English Transformational Grammar*. Waltham: Blaisdell.

Chomsky, N. (1986). *Knowledge of language: its nature, origin and use*. New York: Praeger Publishers.

Chomsky, N. (2001). *Beyond explanational adequacy (MIT Occasional Papers in Linguistics 20)*. Cambridge, MA: MIT Working Papers in Linguistics.

Chomsky, N., M. Halle & F. Lukoff (1956). “On accent and juncture in English”. In *For Roman Jakobson: essays on the occasion of his sixtieth birthday, 11 October, 1956*. Compiled by M. Halle et al. The Hague: Mouton. 65-80.

Chomsky, N. & M. Halle (1968). *The sound pattern of English*. Cambridge, MA: MIT Press.

Chung, S. (1983). “Transderivational constraints in Chamorro phonology”. *Language* 59: 35-66.

Clements, G.N. (2000). “In defense of serialism”. *The Linguistic Review* 17: 181-197.

Cohn, A. & J. McCarthy (1994). “Alignment and parallelism in Indonesian phonology”. Ms., Cornell University and University of Massachusetts, Amherst.
[ROA-25-0894, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Cole, J. (1995). “The cycle in phonology”. In J.A. Goldsmith (ed.), *The handbook of phonological theory*. Oxford: Blackwell. 70-113.

Coleman, J. (1995a). “Declarative Lexical phonology”. In J. Durand & F. Katamba (eds.), *Frontiers of phonology: atoms, structures, derivations*. London: Longman. 333-382.

Coleman, J. (1995b). “Declarative syllabification in Tashlhit Berber”. In J. Durand & B. Laks (eds.), *Current trends in phonology: models and methods*. Volume 1. Salford: European Studies Research Institute. 175-216.

Coleman, J. (2000). “Candidate selection”. *Linguistic Review* 17:167-179.

Collie, S. (2003). *The operation of stress rules in a base-driven, stratified lexicon of English*. Undergraduate dissertation, University of Edinburgh.

Collie, S. (2005). “Why we need Stratal Optimality Theory: the case of English word stress preservation”. In G. Chatzidamianos et al (eds.), *Proceedings of the University of Cambridge third postgraduate conference in language research*. Cambridge: Cambridge Institute of Language Research. 255-262.

Collie, S. (2007). “English stress preservation: the case for ‘fake cyclicity’”. Submitted for review to *English Language and Linguistics*.

Côté, M.-H. (2000). *Consonant cluster phonotactics: a perceptual approach*. PhD dissertation, MIT.

Cutler, A. (1980). “Productivity in word formation”. *Papers from the sixteenth regional meeting of the Chicago Linguistics Society*: 45-51.

Cutler, A. (1981). “Degrees of transparency in word formation”. *Canadian Journal of Linguistics* 26: 73-77.

Davis, S. (1988). “Syllable onsets as a factor in stress rules”. *Phonology* 5: 1-19.

Derwing, B. (1990). “Morphology and the mental lexicon: psycholinguistic evidence”. In W. Dressler, H. Luscutzky, O. Pfeiffer & J. Rennison (eds.), *Contemporary morphology*. New York: Mouton de Gruyter.

Dresher, B.E. (1989). Review of M. Halle & J.-R. Vergnaud, *An essay on stress* (Cambridge, MA: MIT Press, 1987). *Phonology* 7: 171-188.

Elbers, L. (1985). “A tip-of-the-tongue experience at age two?” *Journal of Child Language* 12: 353-365.

Elenbaas, N. (1999). *A unified account of binary and ternary stress: considerations from Sentani and Finnish*. PhD dissertation, Utrecht University.

Elenbaas, N. & R. Kager (1999). “Ternary rhythm and the lapse constraint”. *Phonology* 16: 273-329.

Elfner, E. (2007). “Moraic faithfulness: evidence from Blackfoot and English”. Handout of paper presented at the *15th Manchester Phonology Meeting*, University of Manchester, 26th May 2007.

Fabb, N. (1988). “English suffixation is constrained only by selectional restrictions”. *Natural Language and Linguistic Theory* 6: 527-539.

Fidelholtz, J.L. (1967). “English vowel reduction”. Ms., MIT.

Fidelholtz, J.L. (1975). “Word frequency and vowel reduction in English”. *Chicago Linguistic Society* 11: 200-213.

Field, A. (2005). *Discovering statistics using SPSS (and sex, drugs and rock ‘n’ roll)*. London: Sage Publications.

Fleischhacker, H. (2001). “Cluster-dependent epenthesis asymmetries”. *UCLA Working Papers in Linguistics* 7: 71-116.

Frauenfelder, U.H. & R. Schreuder (1992). “Constraining psycholinguistic models of morphological processing and representation: the role of productivity”. In G.E. Booij & J. van Marle (eds.), *Yearbook of morphology 1991*. Dordrecht: Kluwer. 165-183.

Fudge, E. (1984). *English word-stress*. London: George Allen and Unwin.

Garrett, E. (1999). “Minimal words aren’t minimal feet”. *UCLA Working Papers in Linguistics 1 [Papers in Phonology 2]*: 68-105.

Giegerich, H.J. (1985). *Metrical phonology and phonological structure*. Cambridge: CUP.

Giegerich, H.J. (1988). “Strict cyclicity and elsewhere”. *Lingua* 75: 125-134.

Giegerich, H.J. (1992). *English phonology: an introduction*. Cambridge: CUP.

Giegerich, H.J. (1994). “Base-driven stratification: morphological causes and phonological effects of ‘Strict Cyclicity’”. In R. Wiese (ed.), *Recent developments in Lexical Phonology*. Düsseldorf: Heinrich Heine Universität.

Giegerich, H.J. (1999). *Lexical strata in English: morphological causes, phonological effects*. Cambridge: CUP.

Giegerich, H.J. (2001). “Synonymy blocking and the Elsewhere Condition: Lexical Morphology and the speaker”. *Transactions of the Philological Society* 99: 65-98.

Giegerich, H.J. (2005). “Lexicalism and modular overlap in English”. *SKASE Journal of Theoretical Linguistics* 2/2: 43-62.

Gilbers, D. & M. Schreuder (2002). “Language and music in Optimality Theory”. Ms., Rijksuniversiteit Groningen. [ROA-571-0103, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Goldsmith, J. (1976). *Autosegmental Phonology*. PhD dissertation, Department of Linguistics, MIT.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Oxford: Blackwell.

Goldsmith, J. (1993). “Harmonic Phonology”. In J. Goldsmith (ed.), *The last phonological rule: reflections on constraints and derivations in phonology*. Chicago: Chicago University Press. 21-60.

Gordon, M. (2002). “A phonetically-driven account of syllable weight”. *Language* 78: 51-80.

Gussman, E. (1988). Review of K.P. Mohanan, *The theory of Lexical Phonology* (Dordrecht: D. Reidel, 1986). *Journal of Linguistics* 24: 232-239.

Hale, M., M. Kissock & C. Reiss (1998). "Output-Output correspondence in Optimality Theory". *Proceedings of WCCFL XVI*. Stanford: CSLI Publications. 14pp.
[ROA-202-0697, Rutgers Optimality Archive, http://roa.rutgers.edu]

Halle, M. (1973). "Stress rules in English: a new version". *Linguistic Inquiry* 4: 451-64.

Halle, M. & S.K. Keyser (1971). *English stress: its form, its growth, and its role in verse*. New York: Harper & Row.

Halle, M. & K.P. Mohanan (1985). "Segmental phonology of modern English". *Linguistic Inquiry* 16: 57-116.

Halle, M. & R. Vergnaud (1987a). *An essay on stress*. Cambridge, MA: MIT.

Halle, M. & R. Vergnaud (1987b). "Stress and the cycle". *Linguistic Inquiry* 18/1: 45-84.

Halle, M. & M.J. Kenstowicz (1991). "The free element condition and cyclic vs. noncyclic stress". *Linguistic Inquiry* 22: 457-501.

Halle, M. & A. Marantz (1993). "Distributed Morphology and the pieces of inflection". In K. Hale & S.J. Keyser (eds.), *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press. 111-176.

Hammond, M. (1989). "Cyclic secondary stresses in English". *WCCFL* 8: 139-153.

Hammond, M. (1999a). *The phonology of English*. Oxford: OUP.

Hammond, M. (1999b). "Lexical frequency and rhythm". In M. Darnell et al. (eds.), *Functionalism and formalism in linguistics*. Philadelphia: Benjamins. 329-358.

Hammond, M. (2003a). "Frequency, cyclicity, and optimality". Ms., Department of Linguistics, University of Arizona.

Hammond, M. (2003b). "Frequency, cyclicity, and optimality". Slide show, Department of Linguistics, University of Arizona.
[URL: http://www.u.arizona.edu/~hammond/kslides.pdf (11:32, 13th June 2007)]

Hammond, M. (2004). "Frequency, cyclicity, and optimality". *Studies in Phonetics, Phonology and Morphology* 10/3: 349-364.

Hammond, M. (2005). *An English on-line dictionary*. Department of Linguistics, University of Arizona. [URL: http://www.lexicon.arizona.edu/%7Ehammond/newdic.html#stress (13:40, 13th June 2007)]

Harley, H. & R. Noyer (1999). “Distributed Morphology”. *Glot International* 4/4: 3-8.

Harris, J. (1973). “On the ordering of certain phonological rules in Spanish”. In S.R. Anderson & P. Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt Rhinehart and Winston. 59-76.

Harris J. & E. Gussmann (1998). “Final codas: why the west was wrong”. In E. Cyran (ed.), *Structure and interpretation: studies in phonology* (PASE Studies and Monographs 4). Lublin: Wydawnictwo Folium.

Hartman, J. (2000). “Dictionary pronunciations: mine or theirs or yours?” *American Speech* 75/3: 250-252.

Hay, J. (2001). “Lexical frequency in morphology: is everything relative?” *Linguistics* 39/6: 1041-1070.

Hay, J. (2003). *Causes and consequences of word structure*. London: Routledge.

Hay, J. & H. Baayan (2002). “Parsing and productivity”. In G.E. Booij & J. van Marle (eds.), *Yearbook of morphology 2001*. Dordrecht: Kluwer. 203-235.

Hay, J. & H. Baayen (2003). “Phonotactics, parsing and Productivity”. *Italian Journal of Linguistics* 15/1: 99-130.

Hayes, B. (1981). *A metrical theory of stress rules*. Bloomington: Indiana University Linguistics Club.

Hayes, B. (1982). “Extrametricality and English stress”. *Linguistic Inquiry* 13/2: 227-276.

Hayes, B. (1984). “The phonology of rhythm in English”. *Linguistic Inquiry* 15: 33-74.

Hayes, B. (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.

Hayes, B. (1998). “Phonetically driven phonology: the role of Optimality Theory and inductive grounding”. In M. Darnell et al. (eds.), *Functionalism and formalism in linguistics*. Volume 1: *General Papers*. Amsterdam: John Benjamins. 243–285.

Hicks, C., R. Kennedy, & M. Hammond, (2000). "Lexical frequency and the rhythm rule in English". Paper presented at the annual meeting of the Linguistic Society of America, Chicago, 6th-9th January.

Honeybone, P. (2000). Review of M. Hammond, *The phonology of English: a Prosodic Optimality Theory approach* (Oxford: OUP, 1999). *Journal of Linguistics* 36: 165-171.

Hurrell, E. (2001). *The morphology and phonology of English noun-verb stress doublets: base-driven lexical stratification, prefixes and nominalisation*. PhD dissertation, Department of English Language, University of Edinburgh.

Hyman, L.M. (1977). "On the nature of linguistic stress". In L.M. Hyman (ed.), *Studies in stress and accent*. Los Angeles: Department of Linguistics, University of Southern California. 37-82.

Idsardi, W.J. (2000). "Clarifying opacity". *The Linguistic Review* 17: 337-350.

Inkelas, S. (1998). "The theoretical status of morphologically conditioned phonology: a case study of dominance effects". In G. Booij & J. van Marle (eds.), *Yearbook of morphology 1997*. Dordrecht: Kluwer. 121-155.

Inkelas, S. & C. Zoll (2005). *Reduplication: doubling in morphology*. Cambridge: CUP.

Jackendoff, R. (1975). "Morphological and semantic regularities in the lexicon". *Language* 51/3: 639-671.

James, W. (1890). *Principles of psychology.* New York: Holt.

Jones, D. (2003). *Cambridge English pronouncing dictionary*. 16th edition. Cambridge: CUP.

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. PhD dissertation, Department of Linguistics, University of Utrecht, The Netherlands.

Kager, R. (1993). "Alternatives to the iambic-trochaic law". *Natural Language and Linguistic Theory* 11: 381-432.

Kager, R. (1995a). "The metrical theory of word stress". In J.A. Goldsmith (ed.), *The handbook of phonological theory*. Oxford: Blackwell. 367-402.

Kager, R. (1995b). "English stress: re-inventing the paradigm". Review of L. Burzio, *Principles of English stress* (Cambridge: CUP, 1994). *GLOT International* 9/10: 19-21.

Kager, R. (1995c). “Consequences of catalexis”. In H. van der Hulst & J. van de Weijer (eds.), *Leiden in Last: HIL Phonology Papers I*. The Hague: Holland Academic Graphics. 269-298.

Kager. R. (1999). *Optimality Theory*. Cambridge: CUP.

Kager, R. (2000). “Stem stress and peak correspondence in Dutch”. In J. Dekkers, F. van der Leeuw & J. van der Weijer (eds.), *Optimality Theory: phonology, syntax, and acquisition*. Oxford: OUP. 121-150.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. PhD dissertation, MIT.

Kaisse, E. & P. Shaw (1985). “On the theory of Lexical Phonology”. *Phonology Yearbook* 2: 1-30.

Katamba, F. (1993). *Morphology*. London: Macmillan Press.

Kaye, J., J. Lowenstamm, J.-R. Vergnaud (1985). “The internal structure of phonological representations: a theory of charm and government”. *Phonology Yearbook* 2: 305-328.

Kaye, J., J. Lowenstamm, J.-R. Vergnaud (1990). “Constituent structure and government in phonology”. *Phonology* 7: 193-231.

Kean, M.-L. (1974). “The strict cycle in phonology”. *Linguistic Inquiry* 5/2: 179-203.

Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, M. (1995). “Cyclic vs. non-cyclic constraint evaluation”. *Phonology* 12: 397-436.

Kenstowicz, M. (1996). “Base-identity and uniform exponence: alternatives to cyclicity”. In J. Durand & B. Laks (eds.), *Current trends in phonology: models and methods*. Volume 1. Salford: European Studies Research Institute. 363-393.

Kenstowicz, M. (2005). “Paradigmatic uniformity and contrast”. In L.J. Downing, T.A. Hall, & R. Raffelsiefen (eds.), *Paradigms in phonological theory*. Oxford: OUP. 145-169.

Kenyon, J.S. & T.A. Knott (1944/1953). *A pronouncing dictionary of American English*. Springfield, MA: G.C. Merriam.

Kiparsky, P. (1971). “Historical linguistics”. In W. Dingwall (ed.), *A survey of linguistic science*. College Park: University of Maryland Linguistics Program. 576-642.

Kiparsky, P. (1973). “Abstractness, opacity, and global rules”. In O. Fujimura (ed.), *Three Dimensions of Linguistic Theory*. Tokyo: TEC Company, Ltd. 57-86.

Kiparsky, P. (1979). “Metrical structure assignment is cyclic”. *Linguistic Inquiry* 10/3: 421-441.

Kiparsky, P. (1982). “Lexical phonology and morphology”. In I. Yang (ed.), *Linguistics in the morning calm*. Seoul: Hanshin. 3-91.

Kiparsky, P. (1985). “Some consequences of Lexical Phonology”. *Phonology Yearbook* 2: 85-138.

Kiparsky, P. (1991). “Catalexis”. Ms., Stanford University and Wissenschaftskolleg zu Berlin.

Kiparsky, P. (1993). “Blocking in nonderived environments”. In S. Hargus & E.M Kaisse (eds.), *Studies in Lexical Phonology*. San Diego: Academic Press. 277-313.

Kiparsky, P. (1998a). “Paradigm effects and opacity”. Ms., Stanford University.

Kiparsky, P. (1998b). “Covert generalization”. In G. Booij, A. Ralli, & S. Scalise (eds.), *Proceedings of the First Mediterranean Conference of Morphology*. Patras: University of Patras.

Kiparsky, P. (2000). “Opacity and cyclicity”. *Linguistic Review* 17: 351-365.

Kiparsky, P. (2003a). “Opacity”. Ms., Stanford University. Draft of 1st May, 2003.

Kiparsky, P. (2003b). “Finnish noun inflection”. In D. Nelson & S. Manninen (eds.), *Generative approaches to Finnic and Slavic Linguistics*. Stanford: CSLI Publications. 109-161.

Kiparsky, P. (2007a). “Description and explanation: English revisited”. Slides of paper given at the 81st annual meeting of the Linguistic Society of America, Anaheim, 5th January 2007.
[URL: http://www.stanford.edu/~kiparsky/Papers/lsa2007.1.pdf (11:25, 13th June 2007)]

Kiparsky, P. (2007b). “Reduplication in Stratal OT”. Ms., Stanford University.
[URL: http://www.stanford.edu/~kiparsky/Papers/reduplication.pdf (11:24, 13th June 2007)]

Kiparsky, P. (forthcoming). “Fenno-Swedish quantity: contrast in Stratal OT”. To appear in B. Vaux (ed.), OUP.
[URL: http://www.stanford.edu/~kiparsky/Papers/helsingfors.new.pdf (13:01, 13th June 2007)]

Kisseberth, C.W. (1970). “On the functional unity of phonological rules”. *Linguistic Inquiry* 1: 291-306.

Kozlowski, L. T. (1977). Effects of distorted auditory and of rhyming cues on retrieval of tip-of-the-tongue words by poets and nonpoets. *Memory & Cognition* 5: 477-481.

Kraska-Szlenk, I. (2007). “Analogy: the relation between lexicon and grammar”. Ms., Warsaw University. Version of January 2007. [ROA-893-0107, Rutgers Optimality Archive, http://roa.rutgers.edu/] [To be published in LINCOM Studies in Theoretical Linguistics 38, Munich: LINCOM Europa.]

Kroch, A. (1989). “Reflexes of grammar in patterns of language change”. *Language Variation and Change* 1: 199-244.

de Lacy, P. (2006). *Markedness: reduction and preservation in phonology*. Cambridge: CUP.

Lee, J.-Y. (1996). *Some aspects of English phonology: an optimality theoretic approach*. Seoul: Hanshin Publishing Co.

Leech, G., P. Rayson & A. Wilson (2001). Alphabetical frequency list of the whole British National Corpus, complete without frequency cut-offs. [URL: http://www.comp.lancs.ac.uk/ucrel/bncfreq/flists.html (11:21, 13th June 2007)]

Liberman, M. (1975). *The intonational system of English*. PhD dissertation, Department of Linguistics, MIT.

Liberman, M. & A. Prince (1977). “On stress and linguistic rhythm”. *Linguistic Inquiry* 8/2: 249-336.

Lieber, R. (1980). *On the organisation of the lexicon*. Bloomington: Indiana University Linguistics Club.

Lin, Y.-H. (1997). “Syllabic and moraic structures in Piro”. *Phonology* 14: 403-436.

Lindblom, B. (1986). “Phonetic universals in vowel systems”. In J. Ohala & J. Jaeger (eds.), *Experimental Phonology*. New York: Academic Press: 13-44.

Łubowicz, A. (2002). “Derived environment effects in Optimality Theory”. *Lingua* 112: 243-280.

van der Lugt, A. (1999). *From speech to words*. PhD dissertation, Katholieke Universiteit Nijmegen.

Marantz, A. (1997). “No escape from syntax: don’t try morphological analysis in the privacy of your own lexicon”. *UPenn Working Papers in Linguistics* 4/2: 201-225.

Marantz, A. (2001). *Words*. Ms. MIT. [URL: http://wwwlot.let.uu.nl/zs2001/papersMarantz/WCCFL.doc (20th June 2007)]

Marchand, H. (1969). *The categories and types of present-day English word-formation*. Munich: C.H. Beck’sche Verlagsbuchhandlung.

Marr, D. (1982). *Vision: a computational approach*. San Francisco: Freeman & Co.

Marvin, T. (2002). *Topics in the stress and syntax of words*. PhD dissertation, Department of Linguistics and Philosophy, MIT.

Mascaró, J. (1976). *Catalan phonology and the phonological cycle*. PhD dissertation, MIT.

McCarthy, J.J. (1986). “OCP effects: gemination and antigemination”. *Linguistic Inquiry* 17: 207-263.

McCarthy, J.J. (1998). “Sympathy and phonological opacity”. Ms., University of Massachusetts, Amherst. [ROA-252-0398, Rutgers Optimality Archive, http://roa.rutgers.edu/]

McCarthy, J.J. (2003a). “OT constraints are categorical”. *Phonology* 20/1: 75-138.

McCarthy, J.J. (2003b). “Sympathy, cumulativity, and the Duke-of-York gambit”. Ms., University of Massachusetts, Amherst. [ROA-315-0499, Rutgers Optimality Archive, http://roa.rutgers.edu/]

McCarthy, J.J. & A. Prince (1993a). “Prosodic morphology 1: constraint interaction and satisfaction”. Ms., University of Massachusetts, Amherst and Rutgers Centre for Cognitive Science (RuCCS) technical report 3. [ROA-482-1201, Rutgers Optimality Archive, http://roa.rutgers.edu]

McCarthy, J.J. & A. Prince (1993b). “Generalised alignment”. In G. Booij & J. van Marle (eds.), *Yearbook of morphology 1993*. Dordrecht: Kluwer. 79-153.

McCarthy, J.J. & A. Prince (1995). “Faithfulness and reduplicative identity”. Ms., University of Massachusetts, Amherst, and Rutgers University. [ROA-60-0000, Rutgers Optimality Archive, http://roa.rutgers.edu]

McCarthy, J.J. & A. Prince (1999). “Faithfulness and identity in prosodic morphology”. Ms., University of Massachusetts, Amherst, and Rutgers University. [ROA-216-0997, Rutgers Optimality Archive, http://roa.rutgers.edu/]

McMahon, A.M.S. (1990). "Vowel shift, free rides and strict cyclicity". *Lingua* 80: 197-225.

McMahon, A.M.S. (2000). *Lexical Phonology and the history of English*. Cambridge: CUP.

McQueen, J. (1998). "Segmentation of continuous speech using phonotactics". *Journal of Memory and Language* 39: 21-46.

McQueen, J. & A. Cutler (1998). "Morphology in word recognition". In A. Spencer & A. Zwicky (eds.), *The handbook of morphology*. Oxford: Blackwell. 406-427.

Mohanan, K.P. (1986). *The theory of Lexical Phonology*. Dordrecht: D. Reidel.

Myers, S. (1987). "Vowel shortening in English". *Natural Language and Linguistic Theory* 5: 485-518.

Nanni, D. (1977). "Stressing words in -*ative*". *Linguistic Inquiry* 8: 752–762.

Orgun, C.O. (1994). "Monotonic cyclicity". Ms., University of California, Berkeley. [ROA-123-0496, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Orgun, C.O. (1996a). *Sign-based morphology and phonology, with special attention to Optimality Theory*. PhD dissertation, University of California, Berkeley. [ROA-124-0496, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Orgun, C.O. (1996b). "Sign-based morphology: a declarative theory of phonology-morphology interleaving". Ms., University of California, Berkeley. [ROA-122-0496, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Paradis, C. & J.-F. Prunet (1991). *The special status of coronals: internal and external evidence* [Phonetics and Phonology 2]. San Diego: Academic Press.

Pater, J. (1995). "On the nonuniformity of weight-to-stress and stress preservation effects in English". Ms., McGill University. [ROA 107-0000, Rutgers Optimality Archive, http://roa.rutgers.edu/]

Pater, J. (2000). "Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints". *Phonology* 17: 237-74.

Pinker, S. (1999). *Words and rules*. London: Phoenix.

Plag, I. (1999). *Morphological Productivity*. Berlin: Mouton de Gruyter.

Potter, B. (1994). "Serial optimality in Mohawk prosody". *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 13: 347-361.

Prince, A. (1990). “Quantitative consequences of rhythmic organization”. In M. Ziolkowski, M. Noske & K. Deaton (eds.), *Papers from the 26th annual regional meeting of the Chicago Linguistic Society*, vol. 2: *Parasession on the syllable in phonetics and phonology*. Chicago: Chicago Linguistic Society. 355-98

Prince, A. & P. Smolensky (2004 [1993]). *Optimality Theory: constraint interaction in generative grammar*. Oxford: Blackwell.

Raffelsiefen, R. (1993). “Relating words: a model of base recognition (Part 1)”. *Linguistic Analysis* 23: 3-159.

Raffelsiefen, R. (1999). “Phonological constraints on English word formation”. In G. Booij & J. van Marle (eds.), *Yearbook of morphology 1998*. Dordrecht: Kluwer. 225-287.

Renouf, A. (1987). “Corpus development”. In J. Sinclair (ed.), *Looking up: an account of the COBUILD project in lexical computing and the development of the Collins COBUILD English language dictionary*. London/Glasgow: Collins/COBUILD.

Rice, C. (1996). “Apparent exceptional penultimate stress in English”. *Nordlyd* 24: 157-167.

Rice, K. (1999a). “Featural markedness in phonology: variation – part I”. *GLOT International* 4(7): 3-6.

Rice, K. (1999b). “Featural markedness in phonology: variation – part II”. *GLOT International* 4(8): 3-7.

Ross, J.R. (1972). “A reanalysis of English word stress (Part I)”. In M.K. Brame (ed.), *Contributions to Generative Phonology*. Austin: University of Texas Press. 229-323.

Rubach, J. (1996). “Shortening and ambisyllabicity in English”. *Phonology* 13: 197-238.

Rubach, J. (1997). “Extrasyllabic consonants in Polish: Derivational Optimality Theory”. In I. Roca (ed.), *Derivations and constraints in phonology*. Oxford: Clarendon Press. 551-581.

Rubach, J. (2000). “Glide and glottal stop insertion in Slavic languages: A DOT analysis”. *Linguistic Inquiry* 31: 271–317.

Rubach, J. (2003). “Polish palatalization in Derivational Optimality Theory”. *Lingua* 113: 197-237.

Rubach, J. (2004). “Derivation in Optimality Theory: a reply to Burzio”. *Linguistic Inquiry* 35: 656-670.

Rubach, J. (2005). “Mid vowel fronting in Ukrainian”. *Phonology* 22/1: 1-36.

Rubin, D.C. (1975). “Within word structure in the tip-of-the-tongue phenomenon”. *Journal of Verbal Learning and Verbal Behavior* 14/4: 392-7.

Saciuk, B. (1969). “The stratal division of the lexicon”. *Papers in Linguistics* 1: 464-532.

Sainz, S. (1988). “A noncyclic analysis of English word stress”. *Working Papers of the Cornell Phonetics Laboratory* 3: 1-82.

Sainz, S. (1992). *A noncyclic approach to the Lexical Phonology of English*. PhD dissertation, Cornell University, Ithaca, New York.

Schane, S. (1975). “Noncyclic English word stress”. In D. Goyvaerts & G. Pullum (eds.), *Essays on the sound pattern of English*. Gent: E Story Scientia. 249-59.

Schane, S. (2003). “Dictionaries and phonologists: English accentuation and stress”. *San Diego Linguistic Papers* 1: 129-209.

Schreuder, R. & R.H. Baayen (1995). “Modelling morphological processing”. In L. Feldman (ed.), *Morphological aspects of language processing*. Hillsdale, N.J.: Erlbaum. 131-154.

Scobbie, J., J. Coleman & S. Bird (1995). “Key aspects of Declarative Phonology”. In J. Durand & B. Laks (eds.), *Current trends in phonology: models and methods*. Volume 2. Salford: European Studies Research Institute. 685-709.

Selkirk, E.O. (1980). “The role of prosodic categories in English word stress”. *Linguistic Inquiry* 11/3: 563-605.

Selkirk, E.O. (1982a). *The syntax of words*. Cambridge, MA: MIT Press.

Selkirk, E.O. (1982b). “The syllable”. In H. van der Hulst & N. Smith (eds.) *The structure of phonological representations (Part III)*. Dordrecht: Foris.

Selkirk, E.O. (1984). *Phonology and syntax: the relation between sound and structure*. London: MIT Press.

Selkirk, E.O. (1996). “The prosodic structure of function words”. In J.L. Morgan & K. Demuth (eds.), *Signal to syntax: bootstrapping from speech to grammar in early acquisition*. Mahwah, New Jersey: Lawrence Erlbaum Associates. 187-213.

Siegel, D. (1974). "Topics in English morphology". PhD dissertation, MIT.

Smolensky, P. (1993). "Harmony, markedness, and phonological activity". Ms., Department of Computer Science, University of Colorado. [ROA-87-000, Rutgers Optimality Archive, http://roa.rutgers.edu]

Spencer, A. (1991). *Morphological theory: an introduction to word structure in generative grammar*. Oxford: Basil Blackwell.

Steriade, D. (1982). "Greek prosodies and the nature of syllabification". PhD dissertation, MIT.

Steriade, D. (1994). "Positional neutralization and the expression of contrast". Ms., UCLA.

Steriade, D. (1997). "Phonetics and phonology: the case of laryngeal neutralization". Ms., UCLA. [URL: http://www.linguistics.ucla.edu/people/steriade/papers/phoneticsinphonology.pdf (11:01, 13th June 2007)]

Strauss, S.L. (1983). "Stress assignment as morphological adjustment in English". *Linguistic Analysis* 11/4: 419-427. .

Szpyra, J. (1989). *The phonology-morphology interface: cycles, levels and words*. London: Routledge.

Taft, M. (1979). "Recognition of affixed words and the word frequency effect". *Memory and Cognition* 7: 263-272.

Taft, M. (1985). "The decoding of words in lexical access: a review of the morphographic approach". In D. Besner, T.G. Waller & G.E. Mackinnon (eds.), *Reading research: advances in theory and practice*. Volume 5. London: Academic Press. 83-123.

Teschner, R.V. & M.S. Whitley (2004). *Pronouncing English: a stress-based approach with CD-ROM*. Washington, D.C.: Georgetown University Press.

Wells, J.C. (2000). *Longman pronunciation dictionary*. Harlow: Pearson Education Limited.

Wilbur, R. (1973). *The phonology of reduplication*. PhD dissertation, University of Illinois at Urbana-Champaign.

Wilson, M. (1987). *MRC psycholinguistic database: machine usable dictionary*. Version 2.00. Informatics Division Science and Engineering Research Council, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon. [URL: http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm (15:19, 14th June 2007)]

Yamada, E. (1998). "Problems with stress checking by well-formedness conditions". *Fukuoka University Review of Literature and Humanities* 30/1: 213-236.

Yamada, E. (2005). "Problems with Optimality Theory analysis of English subsidiary stresses". Ms., Fukuoka University.
[URL: http://muse.hum.fukuoka-u.ac.jp/yamada/fuji/sample-E.pdf (10:51, 13th June 2007)]

Yip, M. (1987). "English vowel epenthesis". *Natural Language and Linguistic Theory* 5: 463-484.

Zamma, H. (2005). "Predicting varieties: partial ordering in English stress assignment". Ms., Kobe City University of Foreign Studies/University College London.
[ROA-712-0205, Rutgers Optimality Archive, http://roa.rutgers.edu/]