



universität
wien

DISSERTATION

Titel der Dissertation

“The history of English vowels. An evolutionary account
of the Great Vowel Shift and related vocalic changes”

Verfasser

mgr Kamil Kaźmierski

angestrebter akademischer Grad

Doktor der Philosophie (Dr. phil.)

Wien, 2013

Studienkennzahl lt. Studienblatt:

A 792 343

Dissertationsgebiet lt. Studienblatt:

English and American Studies

Betreuer:

Univ.-Prof. Mag. Dr. Nikolaus Ritt

CONTENTS

<i>List of abbreviations and symbols</i>	v
1 Introduction	1
1.1 Research question	1
1.2 Theoretical framework.....	4
1.3 General hypothesis.....	5
1.4 Assumptions in need of verification and specific hypotheses	7
1.5 Goals	8
2 The theoretical framework	9
2.1 Introduction.....	9
2.2 Evolutionary linguistics	9
2.3 Optimality Theory.....	33
2.3.1 Introduction	33
2.3.2 The architecture of Optimality Theory.....	34
2.4 Evaluation of approaches to language change (and vowel shifting) in the history of English phonology.....	52
2.4.1 Introduction	52
2.4.2 The Neogrammarians	53
2.4.3 The structuralist/functionalist approach	59
2.4.4 Generative phonology	65
2.4.5 OT.....	68
2.4.6 Exemplar-based phonology.....	76

2.4.7	Summary	80
2.5	Formalizing Evolutionary linguistics in OT	82
2.5.1	Some modifications to mainstream OT stemming from its combination with evolutionary linguistics	84
2.5.2	Further advantages to OT	92
3	The relevant developments and a tentative account.....	95
3.1	An overview of vowel chain shifts in English	95
3.1.1	The Great Vowel Shift	95
3.1.2	The Short Vowel Shift.....	102
3.1.3	Contemporary vowel shifts	103
3.2	Study 1: English has become a vowel shifting language	104
3.3	The co-evolution of rhythm and vowels	111
3.3.1	The reduction and loss of unstressed vowels	111
3.3.2	Rhythm-based typology: on stress-timing and syllable-timing	114
3.3.3	Typological perspective	118
3.3.4	A unified perspective	121
3.4	Exaptation of qualitative differences	123
4	The proposed account.....	135
4.1	Introduction.....	135
4.2	Old English	136
4.2.1	Stress assignment in OE.....	138
4.2.2	Fixed lexical stress as a factor enhancing the stress-timing of English through enabling the accumulation of segmental effects of stress	142
4.2.3	Fixed lexical stress as a factor undermining the stress-timing of English though causing stress clashes.....	152

4.3	Middle English.....	157
4.3.1	Stress assignment in ME – grammatical stress	162
4.3.2	Avoidance of stress clashes and lapses	163
4.3.3	Stress assignment in ME – lexical stress.....	165
4.3.4	Summary	167
4.4	Early Modern English	167
4.4.1	The replacement of length with tenseness.....	168
4.4.2	The setting off of the GVS	171
4.5	Present Day English.....	176
4.5.1	Study 2a: Quality is the main expression of vocalic contrasts in PDE	177
4.5.2	Study 2b: Qualitative variation is not predictable from durational variation in PDE	185
4.5.3	A brief cross-varietal comparison	198
4.6	Summary	199
5	Conclusion.....	201
5.1	Summary	201
5.2	Taking stock of a few general insights	203
5.3	Some open questions.....	205
6	List of references	209
7	Appendix 1: Comparison to other languages	230
7.1.1	Romance.....	230
7.1.2	Icelandic	232
8	Appendix 2: A table of vocalic changes.....	235
9	Appendix 3: Materials used for recordings for the study described in Sections 4.5.1 and 4.5.2.....	239

10	Appendix 4: Abstract	244
11	Appendix 5: Kurzfassung	245
12	Appendix 6: Curriculum Vitae	246

LIST OF ABBREVIATIONS AND SYMBOLS

EModE	early Modern English
GMC	Germanic
GVS	the Great Vowel Shift
ME	Middle English
μ	mora
OE	Old English
OT	Optimality Theory
PC	Preserve Contrast
PDE	Present-day English
PIE	Proto Indo European
σ	syllable

1 INTRODUCTION

1.1 RESEARCH QUESTION

The general research question that this thesis sets out to answer is this: why has English become a vowel shifting language? Naturally, lurking in such a question lies already the assumption that it has. Looking at the history of English, one indeed gets the impression that its vocalic system, to some extent stable for long periods of time spanning the separation of English from Germanic up to Middle English times, entered a state of flux, and that this ‘shiftiness’ has been characteristic of it ever since. A major development, or set of developments, that make this transition particularly visible is the Great Vowel Shift (henceforth GVS), through which the entire long vowel system of English was reorganized, in that all long vowels raised by one degree of height, with the exception of the already high vowels /i:/ and /u:/, which diphthongized. If the pronunciations of the following words in the seventeenth century are considered, for instance, then each word had a vowel quality which its neighbor to the right had in the fifteenth century: *make* > *meat* > *meet* > *might*; *boat* > *moon* > *house*. Since the GVS, a number of similar events have been found to take place, including the Short Vowel Shift in early Modern English (lowering and centralizing all short vowels apart from the already low and central /a/), as well as a number of contemporary developments such as the Northern Cities Chain Shift, the Southern Shift, the New Zealand Shift, or the Australian Shift. Given the number of contemporary vowel chain shifts, together with the relative stability of English vocalic system before the GVS, it does seem that English has acquired the trait of ‘shiftiness’, or, at the very least, that the rate at which it undergoes shifts has increased.

The GVS, a major reorganization which affected all long vowels of the language, has been the topic of heated discussions for over a century, proving to be a fertile ground for countless publications. It, and other vowel chain shifts, has been investigated from a number of distinct, if intimately related, vantage points. The present account, in contrast to many

previous attempts, does not seek to explain any of the particular chain shifts, but rather addresses the question of why such shifts are so frequent in Modern English in general.

The first kinds of questions that historical linguists ask themselves are those of the identification and the dating of individual changes. Based on written evidence, they try to reconstruct the relevant developments. Already the first descriptions of it, however, involved attempts not only at reconstructing, but at explaining the mechanism behind it. As soon as the shift was presented in a way which enabled its description to a high degree of generality, that is in a clear pattern, whose graphic visualization is very suggestive, questions as to the causality involved in it began to pose themselves. Thus, the second kind of questions were asked, namely how did the GVS develop, by means of what kinds of mechanisms. An instantiation of this is the controversy whether the GVS was a push chain (Luick 1921-1940) or a drag chain (Jespersen 1909-49), for example, where the descriptions of which vowels moved where are largely agreed on, and the group of events is recognized as a coherent event, but the two views contrast as to what the underlying mechanism was, that is did the vowels concerned cause a shift by vacating a place (a drag chain) or by displacing another vowel (a push chain). Regardless of which account is preferred, though, it is assumed by both that there is a causal link between the individual changes of the shift, and so coherence of the group of events is postulated. This assumption, however, has later been called into question, and there has been a long-standing debate whether any sort of internal coherence can be ascribed to GVS at all (Stockwell & Minkova 1988b; Lass 1988; Stockwell & Minkova 1988a; Lass 1992b). This is the third perspective from which questions about GVS have been asked, and this question boils down to the following: was there even such a 'thing' as GVS? In addition to dating and describing the shift, to explaining the mechanisms behind it, and debating whether internal coherence of it can be assumed which would warrant belief in the very existence of GVS, another question of an explanatory nature that has kept many researchers busy is that of why this shift happened. The search for a triggering event, referred to since Weinreich et al. (1968) as the 'actuation problem' has been riddling many a student of English phonology. This question has been answered in wildly differing ways. Some place the reasons outside of language, starting from the War of the Roses (Ellis, referred to in Wolfe 1972: 7), in the adaptation of East Midland speakers to the prestige Central Southern variety (Perkins 1977), in the attempts of the London bourgeoisie to distance itself from

speakers of the Kentish and Essex dialects (by exaggerating (sic!) the pronunciation of the latter) (Leith 2002: 140); others look for causes within the language itself, as in the physiological aspects of strong stress in English (Lotspeich 1921), the Germanic tendency to root initial stress (Lotspeich 1927), in the raising of the high-mid vowels (Luick 1921-1940: 554), in the application of a natural process which laxes high vowels (Stampe 1972), in the optimization of long high vowels under stress-timing (Donegan 1985), others in the interplay of system internal facts and the functional considerations contingent on the speakers (Martinet 1952).

In the present thesis, these questions, though recognized as important in their own right, are replaced with a more general one. The question asked here concerns the apparent historically observable increase in the likelihood of English vowels to undergo change. Rather than asking why a particular change took place, it asks whether factors can be identified that may have together provided a fertile breeding ground for such changes. This can provide a more general understanding of what happened than either trying to pinpoint triggering events for individual changes or abandoning the question of causality altogether. When it is shown that a set of conditions was met for a continuous parade of vocalic shifts to happen, not finding a particular cause of an individual one becomes more palatable. The wide discrepancies among the different attempts at solving the actuation problem, which differ not only in their conclusion but, first and foremost, in their assumptions about where the trigger for the change should even be looked for, might suggest that these questions should be abandoned altogether, and it may well be that discovering a trigger, or a set of triggers, for each individual development might present insurmountable difficulty. However, once a more global mechanism is known, then not knowing the intricate details of its component parts can be foregone with much ease. Also, the question of coherence of chain shifts is cast in new light. Since it is found that preconditions for unconditioned vowel changes were put in place, the importance of whether or not vowels displace one another, be it by dragging or pulling, loses in importance. When, in a certain scenario, vowels are likely to move, then a vowel leaving its place in the design space is likely to ‘cause’ the movement of another vowel. If the conditions for vowels to move are not there, then, first of all, the first vowel is not likely to move, which, vacuously, will not cause further displacements. The testing ground for this holistic explanation will be English historical phonology.

1.2 THEORETICAL FRAMEWORK

This research question is thought to be best pursued when placed within the framework of evolutionary linguistics. This perspective on language, which assumes that language is an evolutionary system, thus consisting of replicators persisting through time by creating copies of themselves, is capable of solving the teleology problem in language change. Just as the advent of the evolutionary paradigm in biology made redundant previously necessary reference to a designer of the observable purposefulness apparent in all domains of the natural world, so is the evolutionary approach able to account for perceived purposefulness or design without recourse to any designer behind it. Many accounts of language change have been accused of being, implicitly or explicitly, teleological, by resting upon ‘goals’ which give direction to language change. If achieving goals were to be a legitimate mechanism for language change, then either linguistic systems or users of those systems would have to be assumed to possess goals which they want to attain as well as means of fulfilling them. This is obviously untenable with reference to linguistic systems, as no-one would claim that these can wish anything. And speakers, though they certainly have wishes and desires, are definitely not able to foresee the possible consequences of whatever changes they might be capable of introducing into their language, which would be a prerequisite of them shaping long-term developments. With regard to the GVS, the reference to the goal of avoiding mergers as a reflection of the communicative needs of the speakers, the way Martinet (1952) did it, for example, would cast speakers in the role of designers, arriving at the chain shift, with its overall apparent pattern as a best response to the danger of losing lexical contrasts. The very fact that the change took some two hundred years to complete, however, points to the untenability of the agentive role of speakers in shaping the change. This situation, i.e. the presence of design without a designer, is strongly analogous to that in biological evolution.

In order to draw on the findings of the research into sound structure, which extends for well over a century, and to be able to make explicit assumptions about the synchronic organization of these patterns in relevant stages in the development of the English phonological system, a formal phonological framework is adopted. Optimality Theory (OT) has been chosen, because it provides a systematic formalism, which enables explicit

presentation of the proposal put forth here. Additionally, there are some obvious similarities between evolutionary linguistics and OT, in that both frameworks involve constraints selecting among variants. Therefore, the formal apparatus of OT might lend itself more easily to the expression of arguments made from the evolutionary standpoint than formalizations of other frameworks might. Furthermore, the OT incorporates functional motivations into grammar. OT can be seen as functional in its original version of Prince and Smolensky (1993 [2002]), i.e. in that it enables encoding functional considerations directly into grammar as constraints, but it is even more functional in its ‘phonetically driven phonology’ strand (cf. Hayes, Kirchner & Steriade 2004), which seeks to make explicit how functional restrictions enter grammars (Hayes 1996), including the role of perception in contrasts (Flemming 2004). Functional motivations for sound structure need to be expressible in an evolutionary account. For those reasons, and despite the differences between the generative and evolutionary conceptions of language and language change that might transpire throughout this thesis, the formalism developed in OT is used here.

1.3 GENERAL HYPOTHESIS

The account presented in this thesis of how English has become a vowel-shifting language follows the following outline. In PDE, phonetic duration of vowels is influenced by a number of factors, including rhythmic constraints (word stress, foot structure), style, the phonotactic context (coda sonority, foot weight), as well as inherent vowel length. It is proposed here that this complex situation arose in the EModE period, and stands in contrast to earlier stages, at which phonetic duration was first and foremost the expression of inherent vowel length. After vowel duration became tied up in all the other factors, the underlying vowel length contrasts were no longer reliably maintained by it. Still, this did not result in wholesale mergers of morphemes kept apart by vowel length contrasts, which means that phonemes maintaining these contrasts also retained their identity. They did so by ‘exapting’ the qualitative differences that always accompany vowel length contrasts, thus raising these to the status of the primary feature maintaining the phonological oppositions in question. The maintenance of a contrast, though not necessarily by the same feature, can be captured in OT by means of PC (PRESERVECONTRAST) constraints. Under the constructivist view of constraints applied here, this amounts to positing that speakers constructed PC constraints

against leveling out contrasts, but, crucially, since PC constraints do not target specific features, they allowed for the construction of constraints with a different featural content compared to the grammars generating the ambient language.¹ Since these accompanying qualitative differences are such that phonologically long vowels are higher than phonologically short vowels, the exaptation of the qualitative differences, exaggerating them, resulted in the raising of long vowels, and the falling of short vowels. The raising of long vowels (the GVS) and the lowering of short vowels (the Short Vowel Shift), each have the appearance as coherent events; first, since the vowels involved seem to be moving in the same direction (with the exception of the ‘extreme’ vowels in each case, i.e. the already high long vowels /i/ and /u/, which diphthongized, and the already low short vowel /a/, which raised and fronted), and second, since it is possible to conceptualize the movement of one vowel causing the movement of another simply by increasing or decreasing the acoustic space available to the exemplars of the neighboring vowel. The process of one vowel edging out another or ‘attracting’ it by enlarging the acoustic space available to it can extend over long stretches of time, spanning generations, since the systemic pressures are counteracted by the need to imitate others as closely as possible. It is, at the same time, possible for such a systemic pressure to persist nonetheless, since it is inherent in the replicators and the relationships between them, which are transmitted from speaker to speaker through generations. At the time of the exaptation, the relationship between length-induced quantitative variation and the concomitant qualitative variation must have still been present, otherwise the long vowel phonemes would not have mapped onto the tense vowels, and the short vowel phonemes would not have mapped on the lax vowels with such consistency. Afterwards, however, due to the multiplication of contextual factors influencing duration, the co-variation between duration and quality has become much less transparent. For example, the OE /i:/ of *rīden* ‘ride’ pr.subj.pl was longer and higher than /i/ of *riden* ‘ride’ pa.part because of inherent length, and such durational differences were presumably accompanied by qualitative differences, with the /i:/ of *rīden* ‘ride’ pr.subj.pl being higher than /i/ of *riden* ‘ride’ pa.part. One of the contextual factors influencing vowel duration in PDE, on the other

¹ This use of PC constraints does not assign their primary importance in preserving contrasts in input-output mappings, as originally done by Łubowicz, but in modeling speakers’ ability to recognize contrasts, and to attach value to contrasts without necessarily committing themselves to a particular manifestation of these contrasts, during language acquisition.

hand, accounts for the situation in which, for example, the /i:/ in *bead* is longer than the /i:/ in *beat*, without, the former being higher than the latter. Consequently, it is plausible that much more attention is paid to the exact quality of the vowels speakers encounter. Since humans are predisposed to notice and to assign significance to patterns, they might read social significance into the specific vowel qualities, in the face of a lacking clear systemic motivation. As these conditions still obtain, that is there is substantial variation in quality, which is only in part correlated with variation in duration, vowels keep shifting.

1.4 ASSUMPTIONS IN NEED OF VERIFICATION AND SPECIFIC HYPOTHESES

Naturally, the research question, as it stands, is not uncontroversial in itself. It harbors a number of assumptions. First, the question assumes that the degree to which English undergoes diachronic vowel shifts has increased. This assumption begs the question of how this could be measured. Entire vowel shifts, that is coherent events spanning a number of vocalic changes, are not the kinds of entities, as already mentioned, whose existence would be indubitable. And even if it were, they are too big a unit to enable counting across the stages of the development of a language. To answer this problem, vowel shifts are broken down here into their component parts. The sorts of vocalic changes which, taken together, may constitute a vowel shift, are unconditioned changes which do not lead to mergers. A vowel change which affected a given vowel only in a specified environment would not lead to the sort of reorganization that is known as a vowel shift. If only a subset of allophones of a sound moved, and the phoneme remained largely in place in all other environments, than the result would not be wholesale change. That only changes not resulting in merger should be considered follows from the very definition of a shift.

Suggesting a causal mechanism of the type presented here invites questions about its generalizability. The assumption that English has started to undergo more vowel changes as it has become more stress-timed, as well as the assumption that it happened through the loss of predictability of qualitative variation make some typological predictions.

In general terms, one typological implication of the mechanism is that stress-timed languages undergo more vowel shifts than syllable-timed languages. A comprehensive

typological study seeking to ascertain that would go beyond the scope of this thesis, not least because it has been shown that formulating a rhythm based typology might not be feasible (Auer 1993). Still, in order to see how it fares, a brief comparison of the scenario provided here for English with the developments in a number of other languages, namely in Romance, as well as in Icelandic, is provided in Appendix 1.

1.5 GOALS

This thesis strives to present an application of a strictly Darwinian approach to language change in general, and to vowel chain shifts specifically, hoping to show what can be gained by trying to adopt an approach that is as strictly Darwinian as possible rather than one that is merely biologically inspired in a loose way. In particular, the usefulness of the application of the concept of ‘exaptation’, together with all the conceptual apparatus of its source domain, is explored.

In doing so, an overview of major vocalic changes in the history of English is compiled, which might be of interest to anyone who wants an overview of the wealth of the developments in English. As such it is a good entry point into the realm of English vocalic changes.

In addressing the issue of the role of length as a distinctive feature in English it presents a number of arguments, backed up both by a study reported on here as well as by others, and makes a case that length has ceased to be a distinctive feature in English in favor of another feature, traditionally referred to as the tense/lax contrast.

2 THE THEORETICAL FRAMEWORK

2.1 INTRODUCTION

This chapter starts by presenting the general approach to language followed in this thesis, which is that of evolutionary linguistics (Section 2.2). Then, it sketches the main tenets of the formal framework used to formalize the claims put forward in this thesis, namely Optimality Theory (Section 2.3). A selection of previous approaches to vowel shifting is evaluated in light of evolutionary linguistics in Section 2.4. Bringing together OT, evolutionary linguistics and the insights gained from the overview of approaches to phonological change, Section 2.5 outlines the theoretical framework for the discussion which follows.

The theoretical framework outlined in the following is that of evolutionary linguistics, stemming from generalized Darwinism. This general set of assumptions about the nature of language that it presupposes is then formalized with reference to Optimality Theory (OT). A fusion of the two, that is of evolutionary linguistics and OT, is suggested as a way of providing an enlightening perspective on the development of the English vowel system. It is believed that the combined approach has numerous advantages over competing approaches to modeling language change. It can tackle the difficulty (cf. Lass 1980 [2009], 1997) faced by functionalist approaches, which assign the central role to the speakers in the process of language change. By seeing the speakers as one of many environmental factors for the evolution of a language, evolutionary linguistics can factor in their role in language change without committing the fallacy of treating speakers as rational agents behind it. The combined approach also enriches the generative approach to language change, which, by assuming a conceptualization of competence far removed from the psychological reality, runs the risk of postulating teleology in historical developments.

2.2 EVOLUTIONARY LINGUISTICS

The approach to language employed in the present thesis is that of evolutionary (specifically: memetic) linguistics. Memetics is a research program engendered by Dawkins' (1976 [2006]) proposal that cultural evolution can be thought of as strongly analogous to biological

evolution, and his suggestion that a unit of cultural evolution is a ‘meme’, in analogy to the biological gene. Following the publication of Dawkins’ book, the concept caught on, and various attempts to set up memetics as a science of memes were made. Memetics has gained something of a bad name due to various failed attempts at applying it to cultural evolution at large. It is believed, however, that regardless of the lack of spectacular results in the area of culture at large, a memetically inspired, evolutionary account of linguistic, and specifically phonological, change is a feasible project. A rigidly formulated proposition to apply it to model linguistic competence and linguistic evolution is laid out in Ritt (2004). This version of evolutionary linguistic thinking, with strong analogies to evolution in other scientific domains, is presented in more detail after it has been contextualized within the larger picture of different attempts to capitalize on the notion that language is an evolutionary system.

As laid out in Croft (2008), and clear on the perusal of numerous recent publications not mentioned therein, such as Hurford (2012b; Hurford 2012a), Kirby (2007) Smith (2011) (and many others by the members of the Language Evolution and Computation research unit based in Edinburgh <http://www.lel.ed.ac.uk/lec/>), Nowak (2006), or Fitch (2010), both quantitative methods and qualitative concepts of evolutionary science have recently made their way into the study of language. First of all, quantitative methods of evolutionary biology have been borrowed into linguistics, where they have been applied to investigating unresolved issues in linguistic phylogeny, that is the classification of languages into families, (e.g. McMahon & McMahon 2003, 2005). Admittedly, the similarities between the relationships in which languages are to each other on the one hand, and the relationships in which biological species stand to each other on the other hand were recognized much earlier. The realization that languages can be grouped according to the degree of commonly inherited traits goes as far back as the study of evolution itself. Testament to this is the omnipresent representation of language families as branching trees, which developed at the same time as Darwin included his diagram in *On the origin of species* (1859: 17), with some of the first to notice the striking similarities between speciation and the development of languages being Rasmus Rask (1818) and August Schleicher (1863 [1873]). Schleicher’s particular view on the nature of the relationship between biology and languages was largely mistaken, as he saw languages as the analogues of organisms, but the similarities between biological evolution and the development of languages have continued to interest researchers, who have wanted to

exploit these similarities to the benefit of linguistics. For example, long after Schleicher, advanced mathematical methods of modeling phylogeny have been developed in biology, which allow a much more sophisticated reconstruction of speciation and spread of biological organisms (e.g. the Bayes Phylogenies model of Pagel & Meade 2004, 2005). Linguistics has now begun to take over these methods to reconstruct histories of languages (for an overview, see the two sources by McMahon & McMahon cited above), though evading the issue of the nature of the similarity between biology and linguistics. Applying biologically-informed phylogenetic methods to reconstructing language histories can also be helpful in reconstructing human prehistory, as evidenced by the study by Gray et al. (2009), who were able to find support for one of the two competing theories concerning the origin of the settlers of the Pacific².

Aside from borrowing quantitative methods from biology, also computational methods of Artificial Intelligence research have been applied to answering questions of direct relevance to linguistics (for an overview of numerous such applications see Kirby 2002). Methods developed in Artificial Intelligence research have been primarily used with relation to language to probe questions relating to the emergence of language in the first place, that is the transition in the history of humans from the stage with no language to the stage with language. Acknowledging that language evolved at the intersection of biological and cultural evolution, studies employing computer simulations and mathematical modeling have been conducted to test the effects of various mechanisms thought to be important in that process. And so, it has been shown that language may have acquired some of its defining properties (such as syntactic compositionality) simply by virtue of being passed down by a process of cultural transmission. A series of studies employing the so-called iterated learning model, a mode of transmission when one individual acquires a certain communicative behavior by observing another individual who acquired it in the same way have been designed to test that. What they found is that the communicative system becomes more learnable over time. Specifically, this happens by the introduction of compositionality (Smith, Brighton & Kirby 2003, 2005) and of amplification of weak learning biases (Kirby, Dowman & Griffiths 2007). The limitations of those studies, however, become apparent once one notices that, aside from

² Using Bayesian phylogenetic methods Gray et al. (2009) constructed a phylogeny of 400 languages, which supports the ‘pulse-pause’ hypothesis, placing the Austronesian settlers of the Pacific in Taiwan, rather than the ‘slow-boat’ population expansion from Wallacea.

becoming more learnable, the artificial languages used in these simulations become less expressive with each iterative transmission as well, that is the number of distinctions these systems are able to make decreases over time. Thus, there clearly is a mismatch between the situation which obtains in the realm of the simulated artificial languages and natural language, since the latter does maintain a substantial number of contrasts, with this feature being its essential component. Addressing the challenge posed by these findings, laboratory methods with human participants have been developed (for a review of recent developments in this field see: Scott-Phillips & Kirby 2010). The undesired result of the loss of expressiveness can be offset by filtering out homonymy. It can be speculated that a pressure against homonymy is interaction. In fact, comparing mere vertical transmission, where an individual learns a language from an individual who learned it from another individual in a chain of events involving no further communication between the participants, to a situation where repeated pair-wise interaction does take place, have shown that in the latter scenario the expressivity of languages is better preserved (Garrod et al. 2010). Consequently, a strong argument is made to the effect that yes, “repeated individual-level behaviours result in population-level linguistic phenomena” (Scott-Phillips & Kirby 2010), but that they only do so when linguistic behavior at the level of entire populations of speakers is factored in.

Computational studies were also conducted to investigate the role of marking group affiliation for the development and maintenance of language diversity. The results of Nettle and Dunbar’s (1997) simulations suggest that language diversity can be accounted for in evolutionary terms. Cooperation is one of the key characteristics of human societies and its evolution must entail some mechanism of recognizing people who are likely to return the favor from so-called free riders, that is people who accept favors but do not return them. Nettle and Dunbar (1997) argue that dialectal diversity can act well as a safeguard against such free riders. These findings were corroborated by experimental studies with human subjects (Roberts 2008).

While all these studies are primarily interested in the question of how human language emerged and came to have the properties that it has, their findings are also relevant to language change in historical times, since the cognitive apparatus of modern humans must share similarities with that of our predecessors.

Appreciating the achievements of these strands of evolutionary linguistics, which apply quantitative methods developed in other fields dealing with evolutionary systems to questions of linguistic phylogeny and the origin of language, this thesis, however, first of all takes advantage of another link between evolutionary science and linguistics, namely of the applicability of the same qualitative concepts in both domains. Here, two distinct ways of attempting to make use of concepts originating in evolutionary science by the members of the linguistic community can be discerned (Croft 2008: 220). To start with, there have been efforts which recognize the parallels between language and other evolving systems, but which see this similarity as only superficially analogous. Thus, they borrow single concepts formulated by evolutionary scientists without really treating language as an evolutionary system. A case in point is Blevins (2006), whose *Evolutionary phonology* is evolutionary only in the very loose sense of the term, in that it argues that the properties of languages are consequences of their histories. In fact, she does not take the evolutionary perspective much further than already Baudouin de Courtenay (1895) did, as he observed that all alternations seen in languages ultimately derive from historical developments. On the face of it, Blevins' approach is explicitly evolutionary since it employs such notions as inheritance, imperfect replication and natural selection. At the same time, however, the relationship between biology and language is treated as purely metaphorical. Consequently, her model of phonology borrows biologically inspired labels from evolutionary biology for a model that could essentially remain unchanged without them. As a result, the only benefit of even postulating the metaphorical link is the borrowing of biologically-inspired labels into linguistics. Another example of this consequence of seeing the parallels between language and biology as loose and superficial is the fate of the concept of 'exaptation' in historical linguistics, which is discussed in more detail in Section 3.4. In both cases, isolated concepts from evolutionary biology are taken over without the entire conceptual apparatus of the source domain, and so the relationships between the concepts are lost, and potential subsidiary benefits of importing the findings of evolutionary science into linguistics cannot be reaped. Croft (2008: 220) points out another problem with this approach, namely that it might overlook the crucial components of an evolutionary process that make the comparison work in the first place (for example, for any process to be called evolutionary, it must involve clearly defined replicating entities), and, conversely, that it might focus on the particular

mechanisms which are domain specific, but which are not necessary features of evolution. For example, the ‘phenotypic’ expressions of linguistic replicators, that is spoken or written texts, do not have to share the incidental characteristics of the phenotypic expressions of genes, namely organisms, such as mortality. An example of a misconception that has accompanied the evolutionary approach to language from the very start is the notion that languages are like organisms, which was suggested by Schleicher (1863 [1873]: 7), who declared, “[d]ie Sprachen sind Naturorganismen, die [...] entstunden, nach bestimmten Gesetzen wuchsen und sich entwickelten und wiederum altern und absterben [...]”. Such obvious confusion can happen only if the questions of the replicating entities and their expressions are not posed. In fact, the regularities concerning the ways in which languages develop can be likened to the level of not of the organism, but of the population in biology, at which the notions of ageing or dying simply do not apply.

To overcome the weaknesses of a superficial exploitation of the analogy between biology and language, a generalized evolutionary theory is needed. Before various approaches to applying generalized Darwinism to language are presented, a discussion of Darwinism as such and of generalized Darwinism is warranted. Darwin’s (1859) main insight is that the change observable in the evolution of species results from the fact that new traits which may accidentally arise in individuals sometimes increase their chances of reproducing in comparison to other individuals, who lack the innovative trait. Yet, change does not happen ‘in order to’ make individuals or species better. Variation among individuals is the normal state of affairs, and it is the inevitable selection of the more successfully reproducing individuals which gives evolutionary change the appearance of working ‘towards the goal’ of generating ever more successfully reproducing organisms. Darwinian conception of evolution stands in contrast to earlier attempts, most notably to that formulated by Lamarck, in two important ways. First of all, Lamarck proposed that change is goal-oriented, in that modifications arise not accidentally, but in direct response to specific environmental constraints on survival and reproduction. Second, for this goal-oriented view of evolution to work, Lamarck had to assume the inheritance of those acquired innovations by future generations. Darwin’s important discovery was that evolution was in fact blind. He saw variation between individuals of the same species not as arising in order to respond to the environment, but as random, a simple fact of life. It is the survival and reproduction of those

individuals who happen to have the characteristics which best fit the current state of the environment, i.e. natural selection, which results in long-term developments which give the impression of striving after the goal to become better adapted. Thus the second tenet of Lamarck's theory is rebutted, since no change within the lifetime of an individual is passed down, but rather individuals born with a felicitous trait are more likely to survive until reproductive age, and so the new trait can spread. Evidence consistent with Darwin's theory has been accumulating ever since the publication of *On the origin of species* (1859), with two important steps in the process having taken place in the twentieth century. The first important development, whose exact significance to evolutionary biology was unclear at first, is the recognition of the work of Gregor Mendel, who showed that traces of parents do not blend in their offspring. It was later shown by Fisher (1930) that lack of blending in inheritance is indeed not a hurdle, but an essential component of Darwinian evolution, and, as a result, the merger of Mendelian genetics with Darwin's theory of natural selection was born under the name of the modern synthesis. It was further validated by the discovery of DNA by Watson and Crick (1953), which provided the answer to the question of how traits are stored in organisms and passed down.

Despite calls to the contrary, Darwinism, including the strand of evolutionary linguistics stemming from generalized Darwinism adopted here, assumes adaptation to be the only causal mechanism of evolutionary change (for a review of proposals ostensibly undermining the adaptive nature of evolution, and their convincing criticism, see: Dennett (1995: Chapter 10). Granted, there are two other factors which have an effect on evolution, namely drift and catastrophic events. The former cannot result in long-term developments with anything like a discernible direction, since with no environmental pressure present, there is no force to impose apparent directionality on change. The latter, though it might play a role, has to be simply reckoned with as a factor rather than undermining the adaptive nature of the evolutionary change itself. Just as the demise of dinosaurs was precipitated by a catastrophic event, with the event itself not being subject to an evolutionary process, so can external events, such as wars or population migrations, have an impact on the fate of languages. Still, the accounts of which species survived and thrived 65 million years ago, as well as of how languages evolve after they have been placed in a different context by population migrations, can remain strictly adaptive ones.

Historically, there have been numerous attempts to apply evolution beyond its original domain. Two such attempts within linguistics have already been mentioned, an early one, by Schleicher (1863 [1873]), and a recent one, by Blevins (2006). Extrapolating the findings of evolutionary biology to other domains had as its proponent Darwin (1871) himself, who suggested that also languages undergo evolution, as well as Darwin's contemporaries, among them Herbert Spencer, who claiming to have discovered a parallel between biology and sociology, defended social inequality as a reflection of the survival of the fittest. Besides the dubious premises of this parallel, it is impossible to derive ethical claims from scientific theories, as argued already by David Hume, and as argued eloquently with regard to the lack of links between evolutionary theory and ethics by Maynard Smith (1992).

In contrast to these rather flawed attempts to generalize Darwinism beyond biology, there have recently been many more rigid attempts to identify other types of evolutionary processes, besides those genetically based. And so, Cavalli-Sforza and Feldman (1973) propose that cultural inheritance, beside biological inheritance, plays a role in shaping phenotypes. A research program of a more mathematical bent, namely research into complexity and how it emerges through the interaction of a system with its environment, has sought to subsume biological evolution under a more general process governing complex adaptive systems. The most prominent in this area are the researchers associated with the Santa Fe Institute (<http://www.santafe.edu/>), including Holland, Kauffman and Gell-Mann, who include human languages, next to prebiotic chemical evolution, biological evolution, the behavior of vertebrate immune systems, individual learning and thinking in animals, human cultural evolution, the global economy, and artificial intelligence as examples of such systems (Gell-Mann 1992: 8, 1994 [2011]). More recently, also cities have been investigated (Bettencourt et al. 2007), yielding some exciting results. Now, universal or generalized Darwinism is the idea that Darwinian evolution has to apply to life in the entire universe, should life be discovered outside Earth (Dawkins 1983) and that it can be found in domains other than biology (Plotkin 1997), whenever a process of random variation and selective retention of replicators applies. Essentially, Darwinian evolution is 'substrate neutral' (Dennett 1995: 50), that is it operates regardless of the character of the entities to which it applies, as long as it involves entities which are true 'replicators', which are entities with

high copying fidelity, longevity, and fecundity, competing for limited resources (Dawkins 1976 [2006]: 12).

As for the relationship between biology and linguistics, Lass (2003) observes that the borrowing of biological terminology as metaphors into linguistics has indeed been widespread. Given that this is the state of affairs, one could ask whether this is at all helpful. Lass seems to suggest three possibilities here. First, a borrowed metaphor can be of no use at all, merely making the linguistic metalanguage more colorful. Second, if a metaphor is good, it could open up new ways of thinking about issues in linguistics by focusing attention on issues hitherto unnoticed. Third, it could prove to be more than a metaphor, if the exact same processes are shown to be at play in language as there are in biology, in which case one could talk about linguistic evolution in the literal, non-metaphorical sense. Which of the three options is correct is of crucial importance when deciding whether linguists should look to biology for useful insights, and if so how, or if they should sever the ties altogether. In order to decide which of the options holds, the metaphor has to be pushed to its limits, and a concrete fleshed out account of language in evolutionary terms has to be put forward. Even though it is Lass' conviction that "we [linguists] are (or ought to be) a biological science, as the 'social sciences' ought to" (Lass 2003: 50), he stresses that the enterprise of putting the evolutionary approach to the test is valuable, no matter whether it gives green light to strengthening the ties between biology and language or severs them altogether (Lass 2003: 60).

To date, there have been at least three attempts to test the usefulness of the biological metaphor, or rather to test whether it is something more than a metaphor by working out its specifics. These attempts have considered the implications of treating language as an evolutionary system seriously, and seek to take this realization to its logical conclusion, and to investigate language accordingly. The first of them comes from Lass (1997: Ch. 6-7) himself. He sketches a proposal based on the concept developed by Eigen in the field of viral evolution, namely that of a quasi-species. He himself, however, writes this attempt off as an "unsatisfactory discussion" (Lass 2003: 54). Aside from Lass, two specific approaches can be distinguished, each following a different version of a generalized theory of evolutionary change. Croft (2000) bases his on the work of Hull (1988; 2001). Hull ascribes an important role to 'interactors', a term he applies to entities at the level of description referred to in

common parlance as ‘organism’. According to him, ‘genes’ and ‘organisms’ should be replaced with ‘replicators’ and ‘interactors’, which on the one hand should be more fitting concepts within biology, and, on the other hand, enable the generalization of evolution to domains other than biology. By assuming no causal relationship between replicators and interactors, such a model can postulate the existence of replicators at various levels, and this is what Hull does when he sees scientific concepts as replicators and scientists as interactors. Applying this strategy to linguistic behavior, Croft treats units of behavior, specifically tokens of linguistic behavior (‘linguemes’) as replicators, and speakers as interactors. This approach, however, is seriously flawed (Ritt 2004: 180). There is an important issue making utterances unfit for the role of replicators. The most serious problem with this approach is that it overlooks the fact that linguistic units are characterized by their structure and that this structure is constructed by speakers/listeners. As pointed out most clearly by de Saussure (1916 [1995]), an utterance is nothing more than a sound wave, or a pattern of ink in case of written language, and does not per se possess the structural properties which emerge when speakers’ minds interact with it. Sounds and ink patterns cannot be claimed to possess such attributes as ‘being a noun’ or ‘being a phoneme’. Linguistic items, of all levels, are constructed by listeners/speakers when confronted with the speech signal, and so what gets replicated is not the utterance, or a property that the utterance possesses, but a property of the mind producing and receiving speech.

Another formulation of a generalized Darwinian approach is that put forward by biologist Richard Dawkins (1976 [2006]). Dawkins is the most famous advocate of gene-selectionism, the view that can be traced back to Fisher and is shared by, among others, Williams, Maynard-Smith and Hamilton (all referred to in Wright 1980), and which postulates that the gene is the primary unit of selection in biology. The issue of what entities evolution operates on has been subject to a long-standing debate. Traditionally, species were regarded to change. However, species cannot really change in the course of their existence; they emerge as a result of the process of change rather than playing an active part in it (Hull 1980: 327). The two remaining candidates (in biology) are genes and organisms, or rather gene and organism lineages. Hull (1980: 327) defines a lineage as “an entity that changes indefinitely through time as a result of replication and interaction”. Thus, lineages, defined for organisms simply as “evolutionary sequences of ancestral organisms and their descendants” (Russel,

Hertz & McMillan 2011: 423), are the entities that change, and Hull's generalized definition allows the extension of the concept to the cultural domain. While Hull, as mentioned above, settles on 'interactor' lineages (Dawkins' 'vehicles'), Dawkins argues for replicators as the unit of evolutionary selection, and so it is lineages of genes that undergo evolution from this point of view.

For biological evolution, Dawkins took over Williams' (1966 [1996]: 25) definition of a gene, according to which a gene in evolutionary theory is "any hereditary information for which there is a favorable or unfavorable selection bias equal to several or many times its rate of endogenous change". Dawkins (1976 [2006]) makes a very strong case that neither entire species nor individual organisms are the right level of description of how evolution operates, and that, instead, evolutionary changes in the traits of organisms are best understood as happening because of the way in which they affect the fitness of the genes that code for them. Additionally, he sketches a conceivable scenario where biological evolution on earth was preceded by an inorganic evolution, which ultimately resulted in the creation of living organisms. Thus, he argues, evolutionary processes are not confined to the realm of biology, and can take hold whenever the necessary conditions are met. He then proposes that the replicator view, that is considering evolution, metaphorically, as if it were happening 'for the good' of the replicating entities, could also be adopted in cultural evolution, and suggests the 'meme' as the cultural equivalent of the biological gene. He does not provide a definition of memes, but only states that they are analogous to genes and provides a list of examples of what they can be.

Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so memes propagate themselves by leaping from brain to brain via a process which, in the broad sense, can be called imitation (Dawkins 1976 [2006]: 192)

Such a formulation leaves open the question of whether memes should be conceptualized as the external manifestations of some properties of human minds, as properties of minds/brains themselves, or as a combination of the two. This happens to be one of the issues dividing linguists who otherwise agree that languages are true evolutionary systems. The already mentioned approach of Croft (2000) sees cultural replicators as the physical manifestations of

some properties of human minds. The main problem with this conceptualization, as discussed above, is that physical manifestations language, i.e. utterances, do not possess the kind of structure which must be involved in the replication process. Thus, the approach championed in this thesis is that of Ritt (2004), in line with Dawkins' (1982: 109) clarification, namely that memes are the properties of human brains.³

As stated earlier, the initial stimulus for the attempts to set up memetics as a scientific discipline was given by Chapter 11 in Dawkins' *Selfish gene* (1976 [2006]). Despite the fact that Dawkins himself backtracked from the feasibility of serious theorizing about culture in memetic terms in *The Extended Phenotype* (1982), the concept of a meme, and the nascent science of memetics based around it, took on a life of its own. And so, a discussion of the feasibility is taken up by Dennett (1995). Two years later, *Journal of Memetics* (<http://cfpm.org/jom-emit/>) was founded. The most comprehensive attempt at formulating a theory of memetics has been Susan Blackmore's book *The meme machine* (1999). Admittedly, memetics as a theory of culture at large, despite initial euphoria has yet to resolve some serious weaknesses (for a critical discussion of open questions in memetics see e.g. (Rose 1998), (Aunger 2000; Distin 2005). Additionally, the intuitive appeal of talking about cultural replicators as 'viruses of the mind' (1996) is at the same time threatening to extend the meme concept to the extent of conceptual emptiness. Notwithstanding the conceptual difficulties with memetics at large as well as its somewhat bad reputation, it has been applied in linguistics in a way which is very rigorous in its handling of an evolutionary, substrate-neutral conceptual toolkit when applied to language, as described in the following.

A concrete application of Dawkin's view of memetic cultural evolution to linguistics can be found in Ritt (2004), who bases his proposal on the plausible assumption that elements of linguistic competence, or linguistic memes, are physically present in the human brain. Under this view, the unit of replication has clear ontological status, and language ceases to be an entity residing in the timeless and ungraspable realm of Popper's World III, and becomes a World I entity, with its existence firmly rooted in physical reality and thus part of the world that can be subjected to scrutiny. Croft's (2000) model, although it holds the same promise, due to the problems with his idea of linguemes described above, does not provide a satisfying

³ There exist also compromise solutions; McCrohon's (to appear) model, for instance is an example of an approach in which both knowledge and its products are seen as replicators.

solution to the question of ontology of language. By focusing on the external, and, granted, more tractable, outputs of the linguistic ability, Croft's (2000) attempt treats the minds that create these outputs as black boxes. Ritt's attempt, the specifics of which will hinge on the developments in neuroscience, must be right in the central tenet, namely that languages are physically present in human brains. Just as Darwin's formulation of biological evolution was essentially correct, despite Darwin's ignorance of the mechanisms of biological heredity, Ritt's theory may likewise turn out to be at least essentially tenable, even if revisions will be necessary as more is known about how information is stored and processed in the brain. Due to these advantages of Ritt's (2004) model of seeing language in strictly evolutionary terms, this approach is believed to best enable the transfer of the gains of evolutionary science into linguistics.

Following Ritt (2004), then, and earlier contributions (Ritt 1995, Ritt 1996, Ritt 1997) language is conceptualized here in strictly Darwinian terms as a population of replicating constituents. A phoneme, specifically, is thought to be a good candidate for the role of a linguistic replicator. It is here conceptualized as "an association between a complex of specific articulatory gestures, a specific articulatory impression as well as with a set of morph-memes⁴ in whose identification [this phoneme] makes a difference" (Ritt 2004: 170). This conceptualization stands in contrast to the classic Saussurean view, which places all elements of langue, including phonemes, in a super-individual social space, and does not commit itself on ontology, and is therefore not very well suited for evolutionary purposes. Viewing phonemes, instead, as present in brains as associations between articulatory gestures, auditory impressions and morphemes in which these phonemes are distinctive, allows for seeing them as replicators.

To recapitulate the discussion so far, Dawkins' (1976 [2006]) and Dennett's (1995) contributions suggest that evolutionary processes are substrate neutral, that is they could in principle, and, do in fact govern phenomena in domains other than biology, whenever the basic conditions of an evolutionary process are met. The crucial ingredients of an evolutionary process are: replicators marked by a trade-off of longevity, fecundity and copying fidelity, i.e. entities replicating with high, but not perfect fidelity and the differential

⁴ 'Morph-memes' are neural implementations of morphemes.

replication of their copies. The existence of replicators and their selective retention result in the rise in frequency of those replicating entities which are better adapted to the current state of the environment, i.e. in evolution.

As suggested by the two afore-mentioned authors, and implemented in actual linguistic analysis by Ritt (2004), this is the case in language. For Ritt, it is linguistic competence constituents, the prime example of which are phonemes, as implemented neurally in the brain, which can count as replicators in the strict sense. If, as argued by Ritt (1996), phonemes are thought to be physically manifested in brains as association patterns among neural cell assemblies, then they can count as items of a certain longevity. Replicators must exist for long enough for selection to start acting on them. Indeed, phonemes can be thought of as having substantial longevity, as they often remain stably ingrained throughout an individual's lifetime and often subsist longer than that. In some cases, they have gone through numerous cycles of replication without major modifications, as is for example the case with numerous consonants (such as /p, t, k, b, d, g, m, n, w/) which have survived from Germanic all the way to Present-Day English or, to give an example from the world of vowels, the Icelandic short /a/ which, there are good reasons to believe, has remained in the same form since at least the twelfth century. Since by language acquisition they manage to inculcate copies of themselves in new brains before they themselves disintegrate, they possess certain fecundity. In view of the fact that children do not choose the kind of linguistic input they are exposed to, a competence constituent such as a phoneme which is expressed as a systematically occurring contrastive sound in the language produced by a care-taker, is likely to place its copy in the child's brain, and so is characterized by fecundity. As regards their copying fidelity, the continued existence of some constituents in unchanged form in the pool is testament to relatively high fidelity, while, at the same time, it is also clear that copying is not always perfect, since a certain amount of copying error is inevitable. Misparsing on the part of the listener,⁵ as well as articulatory factors (Beddor 2009; Lindblom 1990), result in new variants, i.e. in mutations. The last component of an evolutionary process which gives rise to complexity is the selective retention of the replicators. It is driven by the limited availability of resources, e.g. food, in biology. Given

⁵ along the lines of Ohala (1981)

unlimited resources, all forms could survive, and no apparent purposefulness would arise. This is, however, a purely speculative construct, since the resources are always limited. Just as the environmental resources in biological evolution are, as all earthly things, limited, so do the limitedness of brain space, attention and memory exert a selective force on the replicating language constituents.

Following a similar line of argumentation, Ritt (2004: 143) argues that also units of rhythmic organization, namely feet, might be construed as replicators. To argue for the reality of feet in phonological processing, he points to effects such as the shortening of vowels in longer feet as opposed to their occurrence in shorter feet in English. Such effects can be seen as indicating that speakers are equipped with competence constituents for recognizing such units, since otherwise they would not be able to make the adjustments that they do. The assumption that feet are replicators is crucial for the account to follow, since a particular kind of foot, namely the trochee, is argued to have played an important role in the evolution of English vowels.

On the view, then, that true replicators can be identified when it comes to language, language constitutes under this conceptualization a strictly Darwinian system, and its relationship to biology is argued to be not metaphorical. Rather, both domains are subject to the same generalized principles. The two domains are governed by the same general algorithmic processes, which happen to have been discovered in biology first.

One important thing to note about this view of language right from the start is the role that is ascribed to (the competence constituents of) the language system on the one hand, and to speakers on the other hand. Paralleling the kind of reversal of perspectives in biology, where the gene's-eye point of view is argued by some (Fisher (1930), Williams (1966), Hamilton (1972) (Dawkins 1976 [2006], 1982)), to be more productive than the individual's point of view, the evolutionary perspective places competence constituents at the center of the interest, and positions the speakers kind of in the background (Ritt 2004). Naturally, just as no biologist would deny that the interactions between individual organisms are important for the fate of genes, it is not put in question that speakers are important for the development of language. They are in fact of great importance to this process. However, for the most part, they figure not as agents rationally designing and re-designing their language, but as physical

organisms, that is as one, if very important, set of environmental factors for the propagation of the linguistic constituents, that is replicators. In a way, this could be seen as throw-back to times before the advent of sociolinguistics, when starting with the Neogrammarians, who wanted to place the study of language on firm scientific footing, system internal accounts of change were the predominant mode of explanation in linguistics⁶. In fact, evolutionary linguistics, though not speaker-centered, does not deny their role, nor does it seek to invalidate the results of sociolinguistic research. It just re-casts these influences to shed any recourse to agentivity.

Even though linguistics in the nineteenth century was dominated by the Neogrammarians and so by system-internal modes of explanation, not every linguist working at the time shared this point of view, and there were voices arguing against excluding speakers from accounts of change. Such views can be seen particularly with regard to the issue of the regularity of sound change (e.g. (Schuchardt 1885; Bloomfield 1884). While the question of regularity of sound change is dealt with in more detail in Section 2.4.2, it is worth mentioning at this point, since views on this issue have often run parallel to the views on the role of speakers. The link between the regularity of sound change and the role of speakers is best illustrated with M. Bloomfield's (1884: 178) words, who, doubting the law-like nature of Grimm's Law, writes "a consonant which has been changed could at the will of the speaker have remained unchanged". It has to be noted that the will of a single speaker, even assuming that a particularly resilient individual could stick to a pronunciation on the way out in the speech community, is unlikely to halt sound change, since it takes place at the population level. There was no major change in this respect in structuralism, where system-internal accounts have remained dominant. Just as the conviction about the regularity of sound change was taken over by structuralists from the Neogrammarians, so was the conviction that reasons for sound change are system-internal, and that the will of the speakers should not be brought in as a causal factor. This is the unequivocal position of, for instance (Bloomfield 1933). At any rate, with the work of such figures as Milroy or Labov, speakers were re-introduced as important players in the development of language. Important as it is, factoring in speakers has proven largely problematic. For one thing, as argued vehemently

⁶ For a review of the changing approaches to the role of speakers in explanation in linguistics, see: Deumert (2003).

already by behaviorists (Skinner 1957), studying human beliefs and intentions is extremely difficult. More importantly, accounts seeing speakers as rational agents lack constraint when faced with apparently irrational actions (Deumert 2003), have difficulties relating the micro-level actions to macro-level structures (Deumert 2003; Lass 1997), and in many cases tacitly rest on false assumptions about how language change works (Lass 1980 [2009], 1997). All these conceptual hurdles are better tackled when speakers are an environmental factor and not agents.

There are a number of reasons why adopting the evolutionary perspective on language is not only possible but also desirable (Ritt 1995). The clearing up of an ontological mix-up in the area of language acquisition, the possibility of gaining insights from other fields of study where evolutionary processes have been studied for a longer time, a fresh perspective on the questions of actuation, optimization and so-called conspiracies in language change, are all provided by the evolutionary perspective, as sketched out in the following.

First, approaching language change from the perspective of replicating linguistic constituents obviates the need to mix two levels of description when it comes to language acquisition (Ritt 1995: 53). Under other views on language which refer to language acquisition as an important influence on language structure, such as Lighfoot (2007) or Stampe (1979), language, an abstract system, and speakers, physical entities, come into contact in ways which are not satisfactorily described. In fact, such an attempt does not seem to be so much as undertaken. The only conceivable way to deal with the interaction of the human body and mind is through the brain, and so an account so much as aiming at plausibility has to make reference to some conception of how the brain works. The incorporation of neurolinguistic aspects of language acquisition and use into theories of language change, however, has not been carried very far, and one conceptual difficulty hampering it is the dualism between the abstract linguistic system and its physically existing users, which still dominates the discourse.. As a result, two kinds of accounts of language change can ensue. One solution is to leave as much of the bodily stuff as possible and thus remain within the realm of minds and their creations. Alternatively, one can try to incorporate speakers and their physical properties and needs which stem from them as part of what one describes as linguist and ignore the ensuing dualism altogether. The first solution must by necessity miss out on the important fact that languages are in fact instantiated in

actual human beings, have evolved alongside them and thus their properties must be to some extent contingent on the properties of the speakers. The second solution leads to accounts which allow for the interaction between the physical and cognitive constraints stemming from the fact that speakers are biological organisms on the one hand and languages, which are abstract systems, on the other. To do so, they have to assume a peculiar type of speakers. These have to be speakers who are aware of their needs and limitations as speakers and who are thus capable of bringing them to bear when actively molding and shaping languages to whom they stand in a creator – creation relation. Neither side of this coin is particularly plausible. Both full conscious awareness of one's properties and needs as a speaker and full control over the shape of the language one uses are myths stemming from an erroneous, common-sense understanding of consciousness, with a homunculus-like mental entity overseeing the functioning of the mind and exerting centralized power over it (Dennett 1991). Taking the perspective of linguistic constituents alone, the process of language acquisition can be conceptualized in a coherent way as a process by which linguistic constituents manage to replicate themselves and no reference has to be made to speakers as key players in the process. The child acquiring a language is a generator of random neural configurations, out of which those similar to the ones already in the majority will become reinforced. Thus, children are not ingenious problem solvers re-constructing a language as they see fit, but rather generators of neural schemas, whose variation and selection is not unlike the “brute, mechanical, algorithmic climbing” (Dennett 1995: 75) of Darwinian evolution. In addition, they are mind-bodies, and a crucial part of the environment for competing competence constituents, thus specifying some selective pressures. As a result, no dualism is involved (be it explicitly or, as is often the case, implicitly), and yet all the factors that have been discovered by decades of linguistic research can be incorporated.

Second, evolutionary thinking in biology, as well as in other domains such as evolutionary cognitive psychology and artificial intelligence has produced insights which can benefit linguistics (Ritt 1995: 53). For example, concepts such as ‘competition’, ‘selective pressures’ and ‘environment’ have been well-established in those other domains and can be borrowed into linguistics. Assuming that languages are truly evolutionary systems, the use of the same conceptual framework is not only legitimate, but it also makes possible that findings in the other domains will be able to benefit linguistics. The idea of the competition

of alleles, the various gene variants that can occupy the same stretch of the DNA string, in genetics, for example, can be very illuminating when it comes to explaining language change. In linguistics, any instance of language variation can be thought of as competition between forms, akin to alleles, fighting for a certain structural slot. In both domains change means nothing more than a change in frequencies of the competitors over time. When comparing the state of a population at one point in time to the state of the population at a later point in time, the frequency of one of the competitors, the innovative form, increases relative to the frequency of the resident variant. That there is variation is intrinsic to any replicating system due to the inevitability of copying errors, and so the explanation of a change taking place lies in indicating the selective pressures that were at play in deciding that one of the competitors wins out in the end. The selective pressures are a feature of the environment. For genes, the environment consists of other living beings and features of natural environment, including other genes. For linguistic constituents, the environment can be thought of as consisting, on the one hand, of the properties of the speakers (their physical properties but also all manner of information inhabiting their brains, including their social needs), and, on the other hand, of the properties of other linguistic constituents.

Another concept important in biology which bears special importance with regard to the present thesis, and whose relevance for linguistics at large is discussed therein, is the concept of 'exaptation'. It was introduced into linguistics by Lass (1990), and since then it has been applied in many strands of linguistics for quite some time now. Section 3.4 argues that it is useful in historical phonology, too, but only when seen within the context of evolutionary thinking in general, and not merely as a fancy label. In contrast to many other applications of the concept, its application within a strictly Darwinian approach is advocated here. When exaptation is used as a label for the first stage of an adaptation, the changeability of selective pressures over time, it may be helpful in identifying such transitions. It is more fruitful than taking over the name only, forgetting the source domain and arguing about what it should refer to within linguistics. Considering the entire conceptual machinery, in contrast to simply taking over a biologically inspired metaphor, is where the real potential of the evolutionary approach lies, with regard to this concept and otherwise.

In addition to exaptation, whose usefulness in historical phonology might lie in placing phonologization, an established mechanism of how phonetic patterns enter phonology, in

adaptationist thinking, a concept which is adumbrated to be potentially useful is that of evolutionary extinction generally, and evolutionary suicide specifically. A brief suggestion at how it can help throw new light on the issue of schwa loss in Middle English is provided in 3.3.1.

Third, restating problems of explanation of language change in terms of replicating systems throws new light on the notorious question of actuation of language change (Ritt 1995: 54). Change is inherent in the process of language acquisition under this conception, since acquiring a language consists in trying out various configurations and feedback from the variants already present. The range of the different configurations is not truly random, but circumscribed by the nature of input, by the facts of speech articulation and perception, and by the way the brain operates. That the range of variation is not absolutely random is not an indictment of linguistic mutation as non-evolutionary; what is important is that creation of new variants is not goal-oriented. At any rate, due to this nature of the transmission process, variation in the pool of constituents is only to be expected. Since this happens with every generation of learners, competition is always taking place and so change can proceed. Once a system which comes into being in one individual is in some way different to that present in the speech community, the seed of change has been sown. If the new variant is better, either do its characteristics or to the prestige attached to it, it will spread. Granted, this is not yet a definitive answer to the question of actuation. It is not exhaustive for a question of the type ‘Why did the GVS happen at that particular point in time?’, to be answered by stating that it happened at that time, because it was at that time that the copying errors resulting in the new variants have managed to oust the resident variants. Such a re-statement begs a follow-up question, namely: why did those copying errors manage to do so. A full answer to that question takes into account the advantage that the mutant forms had over the resident competitors at that point in time. The advantages could refer to the ease of perception or production (i.e. to the speakers as environment)⁷ or to the advantages of the mutant form over

⁷ The pressures of the ease of perception and production correspond to the markedness constraints of Optimality Theory, see: 2.3.2

the resident form with respect to the other linguistic constituents (when the whole system of linguistic constituents is seen as environment in which the mutant must find its place).⁸

Admittedly, before a change spreads, it must start in an individual speaker. Here, perceptual factors (Ohala 1981) and articulatory factors (Beddor 2009; Lindblom 1990), which concern individuals, lead, as mutations, to enriching the variation in the pool of structures. As a result, a ‘mini sound change’ (Ohala 1974) in an individual speaker/hearer is not language change yet, but rather an expansion in the overall variation. It is only at the level of a population (of competence constituents) that selection starts to give a certain direction to a change. Hence, perceptual and articulatory factors, which play a role in the creation of new forms, do not themselves set off language change. The change is ‘actuated’ as soon as an identifiable number of speakers have converged on a form perceptibly different from an older one, but it does not make sense to try to pinpoint when exactly actuation happened. The attempt to locate either the first instance of perceptual or articulatory innovation or the onslaught of a qualitatively different variant can never be successful. The former implicates the practical limitation that essentially all of language use would have to be recorded in order to trace back a change once it has happened. The former additionally requires the introduction of an essentially arbitrary boundary for a change to be recognized as such.

Fourth, in a similar vein, questions about the seemingly irreconcilable facts of apparent optimization of language systems on the one hand, and clearly sub-optimal features still permeating them on the other can be tackled in a new way by adopting the evolutionary perspective (Ritt 1995: 55). Under an evolutionary theory, the needs of the speakers are only one of many selective forces influencing the replicative success of a certain replicator. A number of factors affecting the ease of replication itself, with no reference to communication are also at play. Consequently, patterns can get copied simply because they copy well, not because they are particularly ‘useful’. One could argue that this results in a system which is easy to acquire and therefore better than a comparable, but harder one. However, since such a more easily learnable system might very well prove to be unable to make distinctions previously available, a conflict of interests arises. Functionalist frameworks are built around

⁸ The pressures exerted by other linguistic constituents correspond to the constraints on the distinctiveness of contrast postulated by Flemming (2004) and Lubowicz (2003, 2011)

the resolution of such conflicts, but they are (a) inexplicit about the ontological status of those goals and seem to be implicitly suggesting that speakers are consciously aware of their communicative needs and (b) leaving out the ‘needs of the replicators’, besides the needs of the speakers. Since the totality of factors influencing replicability is decisive for the replicative success or failure of individual constituents, suboptimality with regard to the communicative needs of hosts is hardly a surprising state of affairs. For example, phonological mergers, which minimize the number of contrasts expressible by the phonology of a language, do occur.

In historical linguistics, the apparent cases of optimizing language change have sometimes been accounted for by assuming that speakers are rational agents actively directing language change. Even though a comprehensive theory of how they do so has not been developed, and so not much discussion of such mechanisms exists, a number of functionally minded linguists make explicit reference to speakers rationally directing change, which leads to increased functionality, i.e. optimization. Problems with such attempts further highlight the fallacy of assigning agentivity to speakers. A well-known example of a functional, speaker-based explanation is that of the loss of intervocalic /s/ in Greek (Lass 1980 [2009]). This sound change did not take place when it would lead to the collapse of the present and the future paradigms, which has been interpreted either as a case of ‘prophylaxis’, that is of speakers avoiding a change in order not to collapse this distinction (Anttila 1989; Campbell 1975, both referred to in Lass 1997) or as a case of ‘repair’ (Bloomfield 1933, referred to in Lass 1997), that is of speakers reversing the change once they have realized that it collapsed the opposition. Both prophylaxis and repair have been shown by Lass (1997: 355) not to be possible mechanisms of change since, among other reasons, the former would require speakers to anticipate developments that are a long time in the making, often longer than individual life spans, and the latter because lost forms, once gone, are not available to be brought back. A similar case can be observed in Anderson’s (1985: 63) take on Kruszewski and de Courtenay’s views on the mechanism of analogical change. Discussing the presence of the unexpected past participle form *gegessen* of the verb *essen* in German, he explains:

When the past participle of German *essen* ‘to eat’, which we would expect to be **gessen* (<**ge-essen*) added an extra instance of the prefix *ge-* to become *gegessen*,

we can attribute this to the fact that the form **gessen*⁹ **did not seem to speakers** to conform to the principle that German past participles are related to their stems by having the prefix *ge-* before the stem. The residue of subtracting *ge-* from **gessen* is simply *-ssen*, which does not seem to be the root; the entire form is thus integrated into the pattern of the language by taking it as the basis of a newly ‘regularized’ participle *gegessen*. [my emphasis]

Here, it is implied that speakers saw a certain form as non-compliant with a certain generalization present in their language. In other words, this interpretation of analogical change, which takes place because speakers are dissatisfied with a particular form, assumes a need on the part of the speakers to maintain regular paradigms. This need to keep paradigms neat must go hand in hand with intimate knowledge of the generalizations governing those paradigms. Leaving aside the question of the plausibility of these two assumptions for a moment, it has to be acknowledged that the mechanisms of prophylaxis and repair do seem more plausible with regard to this sort of analogical change than with regard to long-term historical change. Speakers would not be required to anticipate far-reaching developments, but, depending on one’s view of phonological processing, either discard a certain form created online in favor of another one, before the violating one were even pronounced (prophylaxis) or having produced or heard the violating form could bar them from reoccurring (therapy). However, the assumed rationality of speakers as taking steps to either preclude or repair the unfortunate creation of a morphologically intransparent form, must be switched off for the step glossed over in the above-cited account. Namely, it assumes the speakers to arrive at **gessen* at one point in their weighing of different forms, which they prefer over the fully regular **geessen*. It is only through the deletion of one of the contiguous vowels that an intransparent form arises. German in fact tolerates words in which verbs beginning in *e-* are prefixed with *ge-*, e.g. *geeignet*, *geerbt*, *gerntet*. Despite the morphological transparency, assumed to be crucially important to German speakers in the above-cited account, and compliance with the requirements of phonological structure on German words, speakers are assumed to arrive at a form **gessen*, which causes them to react against it. These alternative bouts of rational, goal-directed behavior with full-on irrationality

⁹ The form *gessen* is not universally ungrammatical in German, as it is attested in Austrian German (Herbert Schendl, p.c.; Nikolaus Ritt, p.c.). It seems, though, that its ungrammaticality must be assumed for the dialectal lineage in which the form *gegessen* arose.

inherent in such an account, call into question the strategy of invoking speakers who ‘do’ something to their language.

Haspelmath (1999: 188) regards the tendency of describing purposeful design behind seemingly purposeful structures as reflecting an anthropocentric bias. Human-made artifacts are unproblematically talked about in terms of the purposes that their features serve, but problems start when a similar mode of explanation is applied to structures which are not purposefully designed. When investigating language change, a population view is much better suited to accommodate the apparently optimizing character of many instances of language change. The changes are ‘functional’ not because they are undertaken in order to optimize language, but because “there are regular relations between the way language is represented in the mind and the way it is processed during speech production and perception” (Pierrehumbert 2002a: 459), and there are relations between the way languages are represented in the brain and the way languages change. As noted above, evolutionary linguistics in the sense employed here does not see speakers as agents directing language change, and thus escapes all the above-mentioned pitfalls involved in doing so.

Fifth, what can be added to those reasons is the fact that the evolutionary perspective is particularly suited to account for long-term ‘conspiracies’ in language history, i. e. for groups of temporally separated developments which seem to complement each other in striving to achieve a common goal. At least one such conspiracy (proposed by Ritt 2012) is of interest here, namely the development of English rhythm. In brief, a number of developments in the history of English seem to have conspired to make English more compliant with binary trochaic rhythm. The reduction in the complexity of consonant clusters and the rise of distinctive vowel quality for each vowel both made English words more adaptable to trochaic utterance rhythm. It is easier to manipulate the duration of vowels than of consonants, the less complex clusters result in more flexible words. When syllables are stretched or compressed to conform to the trochaic template, vowels can maintain their identity only if they differ in quality, and the faithful expression of their durational differences stands in the way of rhythm. These two developments alone, namely reduction of many consonant clusters and the rise of qualitative differences between vowels previously distinguished phonetically by duration, thus seem to have taken place towards a certain goal, namely making English words better able to express the trochaic foot. Since they took place over a long stretch of

time, vastly longer than that of a single generation, no conscious planning on the parts of the speakers could be imagined, even if the notion of speakers as agents of linguistic change were coherent. At this point, suffice it to say that such a development must be puzzling from a speaker-based perspective. After all, it is inconceivable that successive generations of speakers communicated to their posterity their aims to push the linguistic system in a given direction. From a replicator point of view, however, there is nothing unexpected about the fact that design emerges over long stretches of time, being enacted in successive generations of individuals. Just as in biology genes are selected for their ability to cooperate (Dawkins 1982: 240), those linguistic constituents are transmitted better which form successful coalitions with other linguistic constituents. The time frame is largely irrelevant here, since the same pressures are operative in successive generations of speakers, and their influence on the replicators, i.e. on linguistic competence constituents, are cumulative. And so, even if linguistic systems with vocalic length contrasts expressed by duration were fully functional to their users, slightly different ones, namely those that reinforced them with qualitative differences, were in better agreement with the trochaic foot, and so were able to gain the upper hand in the long run.

As mentioned before, to formalize the proposal regarding the evolution of English vowels put forth here, the formalism developed in Optimality Theory is employed. Before that can be done, the architecture of Optimality Theory must be presented.

2.3 OPTIMALITY THEORY

2.3.1 Introduction

It is believed that a synthesis of evolutionary linguistics and Optimality Theory (henceforth: OT) can result in a potent tool for explaining historical change, and that both OT and evolutionary linguistics can profit from being brought together.

Evolutionary linguistics profits from the fusion because it becomes equipped with a tool for formalizing the selective forces exerted on the replication of competence constituents by the environment. The ‘constraints’ of OT (discussed in more detail below) influence the replication of constituents of linguistic competence (OT ‘input forms’) indirectly, by

constraining their expression (in ‘the output’). While ‘faithfulness constraints’ impose limits on dissimilarity between underlying representations and the output, they facilitate the reconstruction of input forms on the basis of the output in language acquisition. Thus, they increase the faithfulness of the replication of input forms over successive generations of learners. The ‘markedness constraints’, on the other hand, which act against marked, or costly, structures, also affect the replication of inputs, by means of placing limits on the extent to which the complexity of input forms is preserved in the output. The ways in which particular languages at particular stages in their histories respond to the pressures to faithfully transmit contrasts (manifested in the faithfulness constraints) and the pressures to avoid costly forms (manifested in the markedness constraints) can be expressed, in OT terms, as ‘constraint rankings’. These are also transmittable, in addition to the input forms.

OT profits from the fusion for two general reasons; first, because evolutionary thinking, by tying language to its physical manifestation in speakers’ brains makes it a historical object, whose development can then be meaningfully studied, and second, because population thinking makes the conceptualization of synchronic variation and historical change much easier. Hence, evolutionary thinking makes OT better adjusted to modeling language change, a task which it was not primarily constructed to perform. Admittedly, there have been many applications of OT to language change (see e.g. contributions in Holt 2003), and the present thesis draws on their findings, but the biologically enriched perspective is believed to provide a more enlightening treatment of the issues discussed therein.

2.3.2 The architecture of Optimality Theory

Optimality Theory is a grammatical framework which is one of the descendants of Classic Generative Grammar. The first comprehensive presentation of OT is Prince and Smolensky’s (1993 [2002]) classic work *Optimality Theory: Constraint interaction in generative grammar*. As laid out therein, the major difference between OT and previous generative approaches to grammar is that OT dispenses with the transformational rules which spell out the relationship between the input and the output, a hallmark of generative phonology since Chomsky and Halle’s (1968 [1997]) *Sound pattern of English* (from now on: *SPE*) and a feature of many of its off-shoots such as Lexical Phonology (Kiparsky 1982b; Giegerich

2005) or Natural Phonology (Stampe 1979), and postulates in their place universal, violable constraints.

In keeping with previous approaches, grammar in OT is seen as an input-output mechanism mapping the relationship between underlying representations and surface forms. This goal is achieved by reference to ordered rules in older frameworks, and by reference to constraints in OT. Two specific problems inherent in rule-based phonology that OT has overcome because it is based on constraints and not on transformational rules are the Duplication Problem (Kenstowicz & Kisseberth 1977: 136) and the issue of Conspiracies (Kisseberth 1970). The Duplication Problem is the observation that rule-based theories posit limits on the underlying structure of morphemes (so-called Morpheme Structure Constraints), which produce morphemes which are identical to outputs of separately posited transformational rules. These two findings must clearly be related, but it is not possible to capture this relatedness by means of rules. What is referred to as a conspiracy by Kisseberth (1970) is the observation that the same goal, say an elimination of consonant clusters, can be attained by separate rules, whose form does not state in any way that they are related. That they are related, however, is beyond doubt, and where their relatedness resides is in their effect. A rule-based theory has no way of capturing this. These two instances of loss of generality are overcome in OT, as should become clear from the following brief sketch of its basic architecture.

In Classic Generative Grammar, the role of phonology consisted in applying a number of ordered transformational rules to lexical representations to arrive at the output forms. With time, the unconstrained nature of the possible scope of the form of re-write rules and their interactions has led both syntacticians and phonologists to postulate universal conditions on rules and rule interactions, culminating in the formulation of the Principles-and-Parameters theory. An example of such a postulated universal condition on phonological rules is the Obligatory Contour Principle of Autosegmental Phonology, which, in its strongest formulation, bans the association of identical tones with sequences of adjacent vowels (Odden 1986). The need for those universal conditions, however, has in turn resulted in the increased abstractness of analyses and the multiplication of levels of derivation. Being universal, the principles could not be violated, and so levels of derivation were postulated at which the universal conditions were actually satisfied. Thus, the goal of constraining rules

and their interactions was actually subverted. In order to rescue the endeavor to constrain phonology by incorporating universal conditions, a different take on the universality of the conditions on rules and their interactions has been suggested. If universality is thought of in terms of markedness, that is relative ill-formedness of some forms with relation to others, then the universal conditions become violable.

The modern understanding of the concept of markedness dates back to Trubetzkoy (1939). Within the context of the discussion of neutralization, he proposes that when a certain binary opposition is neutralized, and when the result of the neutralization is one of the two sounds in question, and when there are no contextual reasons for why one member of the opposition should surface rather than the other, then the result of the neutralization is the ‘unmarked’ member, as opposed to the ‘marked’ one (‘merkmallos’ vs. ‘merkmaltragend’ Trubetzkoy 1939: 73). The idea is that the sound surfacing with no contextual trigger should be considered to require less specification, and it is the other member of the opposition that has to be marked for the feature. Under this understanding, markedness of a segment is language-specific (Rice 2007: 79, 86). Extending Trubetzkoy’s idea, Jakobson (1944 [1969]) availed himself of the concept, and formulated implicational laws, under which marked structures imply unmarked structures in a variety of domains, including development of sound systems in child language, the degradation of sound systems in speech deficits, and diachronic development of languages. These two early applications of markedness seem to form the basis for two broad traditions of invoking it. As laid out in Rice (2007: 80), one could distinguish between phonological markedness (also called ‘structural markedness’ Bybee 2001) on the one hand, and ‘natural markedness’ (Anderson 1985) (also referred to as ‘frequency markedness’ Bybee 2001) on the other hand. Phonological markedness is explicitly postulated to be a property of synchronic grammars, detectable by a number of diagnostics relating to the behavior of sounds in phonological patterning, such as neutralization, epenthesis, assimilation, coalescence or deletion. Natural markedness is essentially a cover term for all the other interpretations of markedness, including implications for language acquisition, typological frequency, ease of articulation, perceptual salience, and diachronic stability.

Phonological markedness, as the ‘Markedness Theory’, was used by Chomsky and Halle (1968 [1997]: Chapter 9) as an extension of the theory presented in *SPE*, to allow for the

discrimination between likely and unlikely grammars. For them, though, markedness was in fact extragrammatical. They propose that features can be ‘marked’ and ‘unmarked’ in addition to ‘+’ and ‘-’, in the feature matrix of lexical entries. These are subject to markedness conventions, which are supposed to be universal. Stampe (1973) noted a number of problems with such treatment of markedness. One of his criticisms is that the proposed universality of markedness conventions cannot deal with cases of (language-specific resolution of) conflict between various such conventions.

In OT, conflict between the contradictory requirements imposed on the output forms by different markedness considerations is resolved by stating the markedness of particular structures in the form of markedness ‘constraints’, and by ranking those constraints with respect to each other for a given language/language state. As far as the technical meaning of markedness in OT is concerned, Kager (1999: 2) states that under markedness, linguistic elements are not universally well or ill-formed, but only better or less-well formed relative to other linguistic elements. Additionally, the markedness of elements is rooted in perceptual or articulatory phonetics. Thus, grammar is constrained since the universal markedness conditions are not arbitrary, and at the same time, no intermediate stages of derivation have to be postulated since markedness conditions are relative and the violation of some of them might be allowed, provided that others are satisfied. Such a conception of universal conditions combined with reference to phonetically grounded markedness finds its expression in the ‘constraints’ of Optimality Theory.

The constraints at the core of OT are of two types; there are markedness constraints and faithfulness constraints. Markedness constraints are restrictions on the well-formedness of the output and faithfulness constraints are restrictions on the disparity between the output and the input forms. The incorporation of markedness considerations into the very heart of grammar, in the form of markedness constraints, is one of the key features of OT. Markedness constraints are grounded phonetically, either articulatorily or perceptually. The markedness of front round vowels, for instance, is grounded in acoustics. Lip rounding has an effect of lowering all formants and since front vowels are characterized by high values of the second formant, the lowering effect of lip rounding obscures the values of the second formant cuing the frontness, and so front and round do not make for a good combination. Thus, a constraint against front rounded vowels, *[+front, +round] is part of the grammar of

every language, and the surfacing of such vowels in some languages is accounted for by this constraint being outranked by a faithfulness constraint militating against the ‘repair’ of the marked, front rounded vowels. Although grounded in phonetics, markedness constraints are still phonological, because (a) they are categorical while phonetics is gradient and (b) the value attached to particular markedness constraints differs from language to language (Kager 1999: 5), and (c) phonological constraints are symmetrical, whereas phonetic difficulty scales are asymmetrical (Hayes 1996). Markedness constraints operate on output forms only, while input forms are completely free of any markedness considerations, which was dubbed by Prince and Smolensky (1993 [2002]: 209) as the principle of ‘the Richness of the Base’. The input forms, incidentally, are also free to display features which do not have to be distinctive, as shown by Kirchner (1995; Kirchner 1997). The curious effect of this approach to the input is that it contains, in a sense, more information than the output. Since the presence of markedness constraints on the output leads to the disappearance of at least some contrasts present in the input, and since the lack of corresponding faithfulness constraints will lead to the disappearance of non-distinctive features (Kirchner 1997), the output will always be poorer in information than the input. The second group of constraints, i.e. the faithfulness constraints are the forces which require the preservation of lexical contrasts. In order to be able to convey contrasts in meaning, any language must be able to produce a certain number of formal contrasts, and the preservation of these is handled by faithfulness constraints. Faithfulness constraints make reference both to input and to output. These two types of constraints are inherently in conflict, since a maximally faithful grammar would be one that allowed all conceivable segments and all their logical combinations, which would result in lexicon entries containing highly marked sound strings, which would also be allowed to be of any length. On the other end of the extreme, a grammar maximally concerned with markedness would allow a very restricted number of segments to surface in the output, and would impose heavy restrictions on their combinations in words. By Kager’s (1999: 6) estimation, the former kind of extreme scenario would allow for the creation of 300 billion potential lexical items, and the latter of only 36; since, according to Kager (1999: 7) an average natural language needs to store some 100,000 items, a compromise between markedness and faithfulness is not only imaginable, but in fact expected.

The conflict between various markedness and faithfulness constraints which stems from their often contradictory requirements is resolved by means of ranking the constraints. Within the grammar of each language, the entire set of constraints forms a hierarchy with each higher ranked constraint strictly dominating all of the constraints below it. Whereas the constraints themselves are universal, and thought to be part of the Universal Grammar (though this view has been relativized, see Section 2.5), the ways they are ranked with respect to each other are language specific and have to be acquired by language learners. The optimality of the output is the degree to which it violates constraints. As already indicated, that idea that the optimal output will not violate any constraints is misguided. This is not the case in view of the fact that constraints are in conflict with each other, since they have contradictory requirements, i.e. satisfaction of some constraints must lead to the violation of others. Consequently, each potential output is bound to violate some constraints. For each language, though, the universal constraints are ranked, and the higher ranked constraint dominates all lower ranked constraints. The optimal output, then, is the most harmonic one, which is the output which least violates the ranking of conflicting constraints. To formulate a grammar of any (stage of a) language, the ranking of relevant constraints has to be formulated, since “[l]anguages basically differ in their ranking of constraints” (Kager 1999: 4).

As stated above, OT, similarly to other generative frameworks, sees grammar as “an *input-output mechanism* that pairs an output form to an input form (such that each input has precisely one output)” (Kager 1999: 18). In OT, this mechanism consists of two components. The first component, the so-called Generator (or Gen) generates an infinite number of possible outputs for a given input. According to the already mentioned tenet of ‘the Richness of the Base’ (Prince & Smolensky 1993 [2002]: 209), i.e. the assumption that no constraints apply at the level of the input, even highly marked structures can be put forward by Gen, and no limits on this generation whatsoever hold at this stage. The assumption that no constraints make exclusive reference to input morphemes overcomes the Duplication Problem inherent in the rule based theory, that is the need to postulate constraints on the structure of morphemes which mirror the outputs of rules, with no recognition of the similarities between the two things. After candidate outputs have been generated by Gen, they are assessed by the second component, the Evaluator (or Eval) with respect to how harmonious they are. Their

harmony depends on how well they conform to the markedness constraints on the one hand and to faithfulness constraints on the other hand, or, seen from the other perspective, how badly they violate these two types of constraints. Eval consists of the language-specific ranking of universal constraints together with a system of assessing their harmony by awarding violation marks. In the end, Eval selects one and only one optimal output form from the entire candidate set submitted by Gen. From the above characterization it is clear that it is Eval that is burdened with doing the bulk of the work of producing the right output, since Gen's task is rather automatic and completely unconstrained. In addition to these two components, Gen and Eval, grammar additionally contains the Lexicon, where the lexical representations of morphemes are stored. Overall, the major task phonology has to fulfill is very different to that known from transformational models of phonology. It is not to produce the output form knowing the input form and a transformation it is supposed to undergo. It is, conversely, to generate a host of candidate output forms, and select the optimal one. This assessment is done on the basis of comparing the output to a language specific ranking of a set of universal constraints.

In classical OT, this evaluation of candidates happens in parallel, so no levels of derivation or ordering of any kind are postulated. This position of mainstream is a reaction against the lack of principled constraint on the ordering of rules in classic Generative Phonology. Before OT, however, there already was an attempt to limit the number of derivational levels allowed by the theory by positing a number of cycles at which all rules applicable at that cycle must apply in parallel. This attempt was Lexical Phonology (Kiparsky 1982b), and there is a continuation of this solution to the ordering problem within OT, known as Stratal OT (Bermúdez-Otero 2007, in preparation) or 'LPM-OT' (Lexical phonology and morphology OT) (Kiparsky 2000). The assumption of this strand of OT that constraint rankings select candidates in a cycle at three successive levels, the stem level, the word level and the post-lexical level, is followed in this thesis.

One of the most influential changes to the basic OT architecture were made by McCarthy and Prince's (1995a; 1999) Correspondence Theory. Wanting to capture the parallels between base/reduplicant relations and input/output relations, McCarthy and Prince propose that a correspondence relation obtains between input forms and output forms, just as between the base and the reduplicant. The correspondence relation is defined as follows:

(1) Correspondence

Given two strings S_1 and S_2 , **correspondence** is a relation \mathfrak{R} from the elements of S_1 to those of S_2 . Elements $\alpha \in S_1$ and $\beta \in S_2$ are referred to as **correspondents** of one another when $\alpha \mathfrak{R} \beta$.

Based on this relation, McCarthy and Prince (1995a) introduce three new families of constraints, namely the MAX constraint family, the DEP constraint family and the very widely used IDENT(F) constraint family. MAX constraints require that a feature present on the primary level (e.g. input, base) be present in its correspondent on the secondary level (in e.g. output, reduplicant). DEP constraints require that a feature present in the secondary level (e.g. output, reduplicant) be present in its correspondent in the primary level (e.g. input base). IDENTITY constraints require that a feature specification present in two correspondents be identical.

Incidentally, the introduction of IDENT and DEP constraints, and their role in processes such as epenthesis and deletion, diminishes the role of a separate, phonetic component, whose other job has been to furnish phonological surface language specific, universally non-contrastive detail. The view of phonetics as having to fill out the language specific detail was criticized e.g. by Pierrehumbert (1994), and it was shown by Kirchner (1997) that the phonetic component can be relieved of this duty by means of faithfulness constraints¹⁰. As a mapping in which an input matches the featural specification of the output will cause fewer violations than a mapping which is identical with respect to all but this input specification, fully specified inputs will prevail¹¹.

This is so despite the fact that, due to Richness of the Base, the lexicon is not constrained by any considerations of markedness. Still, the lexical entries harbored in the brains of actual speakers will in practice be very close to surface representations, and they will even include contextually predictable features, due to a principle called ‘Lexicon Optimization’ (Prince & Smolensky 1993 [2002]: 191). It posits that, when faced with a lack of evidence for a particular input, learners will choose such an input that results in mappings producing the least constraint violations. To illustrate it with an example, the case of aspiration of voiceless

¹⁰ Although Kirchner (1997) does not employ IDENTIO constraints, the constraints he uses can be thought of as equivalent to IDENT constraints.

¹¹ The exact nature of the phonetic component in OT is rarely discussed. One exception is Hayes (1996), who notes that the phonetic component is also likely to be OT-like.

plosives in English can be considered. Aspiration can be said to be a predictable, or redundant feature in English plosives. They are aspirated in prevocalic position in stressed syllables, unless preceded by /s/; elsewhere, they are unaspirated. Now, due to ‘Richness of the Base’, for a word like *cat*, the grammar must select an output candidate with aspiration in the initial plosive regardless of the presence or absence of aspiration in the input. This can be achieved by ranking a context sensitive markedness constraint against unaspirated voiceless stops in the onsets of stressed syllables (let us call it *'[-voice, +stop]V) higher than a faithfulness constraint requiring the identity of the feature [aspiration] between the input and the output (let us call it IDENTIO(aspiration)). If *'[-voice, +stop]V dominates IDENTIO(aspiration), then the initial plosive of *cat* will be aspirated in the output, regardless of whether aspiration is present in the input or not, as illustrated in Tableau 1 below.¹² Regardless of the input, an output candidate without aspiration on the initial plosive will incur a fatal violation of the highly ranked markedness constraint.



Input: /kæt/	*'[-voice, +stop]V	IDENTIO(aspiration)
kæt	* !	
 k ^h æt		*
Input: /k ^h æt/		
 k ^h æt		
kæt	* !	

Tableau 1: Aspiration in voiceless plosives

However, even though such a grammar will yield the correct output regardless of the input, under Lexicon Optimization, a language learner will still posit /k^hæt/ as input. This is so, because the mapping /k^hæt/ → [k^hæt] will not result in a faithfulness violation, whereas the mapping /kæt/ → [k^hæt] will (Steriade 2007: 152).

2.3.2.1 Vowel representations

An issue that does not constitute the primary concern for a lot of research in OT, which however cannot be left out here, is the issue of phonological representations, and of segments

¹² This is a slightly simplified scenario. To capture the full pattern of allophonic distribution, a context-free markedness constraints against aspirated stops would have to be introduced and ranked between the two constraints already present, to make sure that voiceless stops are not aspirated regardless of position, i.e. to rule out forms such as [k^hæt^h]

specifically. Continuing the shift of focus away from representational inventories, which was the preoccupation of structuralism, onto grammars, as seen with the rise of generativism (Anderson 2000), OT is primarily concerned with the linguistic knowledge of speakers that relates the underlying forms to the surface forms. This linguistic knowledge was manifested in transformational rules under Classic Generative Theory, and is postulated to reside in constraint rankings of individual languages in OT. Thus, the form of constraints and the interactions between them, which is the core of grammar for OT, receives more scholarly attention than representations. Additionally, going one step further away from the preoccupation with representations compared with the transformational accounts, OT actively argues for the irrelevance of the input in a range of phenomena, e.g. in explaining allophonic alternations, which follows from the already mentioned tenet of the richness of the base; because even highly marked sequences are allowed at the level of input, the particular constraint ranking must be shown to yield the correct result even when faced with such inputs. Pragmatically speaking, one does not easily find a list of features the OT representations assume in Prince and Smolensky (Prince & Smolensky 1993 [2002]) or in Kager (Kager 1999) and one must contend with the observation that “[m]ost OT literature on phonology [...] assumes the representational alphabet of non-linear (metrical and autosegmental) phonology” (Kager 1999: xii). Since the present contribution is an investigation of a history of the sounds of English, issues pertaining to segmental representations are of some interest for the following presentation. Consequently, the key insights of nonlinear phonology are sketched in the following. Nonlinear phonology arose primarily in response to the difficulties faced by the linear representation of segments as ‘feature bundles’ when confronted with representing phenomena such as tone and stress (McCarthy 1982; Pulleyblank 1989). The classic generative view on the representation of segments, known since Jakobson, Fant and Halle’s (1951 [1963]) theory of distinctive features, but most radically expressed in *SPE*, was that segments amount to nothing more than vertical columns of unordered features in feature matrices of lexical entries. This view had at least two major weaknesses.

The first problem stemming from such a representation has to do with the ‘one feature – one segment’ principle implicit in it. It was observed that certain ‘features’, e.g. tone, can sometimes span more than one segment, or, alternatively, two different values of those

features can be manifest in individual segments. There was no way in which this could be properly accommodated within a framework based on feature matrices. To enable the representation of features that are longer or shorter than segments, Goldsmith's (1976) *Autosegmental Phonology* postulated a view of phonology with separate, additional tiers of representation where these autonomous features reside. Features on different tiers are connected with 'association lines', and a one-to-one correspondence yielding one feature per segment, which used to be the only possibility in the linear representation, is just one of many logical possibilities for this relation under the autosegmental view. This approach to phonology was first demonstrated to better deal with prosodic, or 'suprasegmental' phenomena such as tone, stress, and intonation. However, it was soon shown that the autosegmental view was superior to the linear view also with regard to other features, hitherto treated as segmental, such as [nasal], which may extend over longer stretches of segments or be present in only a portion of a segment; or the vocalic features important for vowel harmony, namely [back], [round] and [ATR], which can span many syllables (Clements & Hume 1995: 246).¹³

The second problem stemming from the linear conception of segments is that it has no way of accounting for the reality of groupings of individual features into such classes as 'place' or 'manner' features in phonological patterning. They have been repeatedly shown to play an important role, and yet they are, under the matrix view, notationally no better than any arbitrary grouping of features. A common example suggesting the reality of some sort of unity of place features is the prevalence of place assimilation phenomena in various languages such as Spanish, where the nasal consonant of the indefinite article *un* assimilates its place of articulation to the following consonant. Rather than having to specify the process for each possible following consonant separately (e.g. /n/ becomes velar before velars, labial before labials and so on), one can state this behavior in general terms, by saying that "nasal consonants share the place node with a following consonant" (Clements 2006). A process assimilating the place of articulation of one of two contiguous consonants to those of the other has to specify each of the individual features separately, which does not distinguish

¹³ Interestingly, the possibility to account for features spanning domains longer than one segment was already known to structuralists, notably Zellig Harris and Charles F. Hockett (Clements 2000), and even a part of Chomsky's pre-SPE publications (Encrevé 2000), but it was eradicated by the radical, linear approach of the *SPE*.

such a process from a whole host of other, unattested processes which would make reference to an arbitrary set of other features (Clements & Hume 1995: 248). Here, again, pre-SPE phonologies recognized this need, e.g. Trubetzkoy's (1939) proposal of 'related features', which was eclipsed with the onslaught of generative phonology. Approaches to phonological representations in the 1980's (Clements 1985; Sagey 1986) which seek to factor in the groupings of features into classes propose a model of internal organization of features now known as 'feature geometry', under which "segments are represented in terms of hierarchically-organized node configurations whose terminal nodes are feature values, and whose intermediate nodes represent constituents" (Clements & Hume 1995: 249). One of the benefits of this representational model is that it is capable of distinguishing between probable and improbable rules or processes in that it restricts the number of conceivable processes by allowing only those which make reference to constituents in the structure and banning those that would make reference to a set of features which did not form a constituent.¹⁴

As for the representation of vowel height, Clements and Hume (1995) draw on Clements (1991) who suggests analyzing vowel height by making reference to a single feature [open], which can be represented at several tiers. At the highest tier, the value of the feature divides vowels into two groups; high and low. At the next level, the feature can again take either a positive or a negative value, to make further divisions within the low and high vowels as assigned at the highest tier. Thus, two tiers enable representing four distinctive vowel heights. If further divisions are necessary, then another tier can be postulated (Clements 1991). This approach is different to the traditional one with two features, namely [\pm high] and [\pm low]. Having those two features enables the representation of only three heights, since the combination [+high, +low] is not possible. To account for more distinctive vowel heights, additional features have been brought to bear. Chomsky and Halle (1968 [1997]) took over [\pm tense] from Jakobson, Fant and Halle (1951 [1963]), and others, generalizing Stewart's (1967, referred to in Laver 1994: 289) proposal for Akan have used [\pm ATR] (advanced

¹⁴ An alternative solution to the problem of distinguishing between probable and improbable rules is that proposed by Natural Phonology (Stampe 1979; Donegan & Stampe 1979; Dziubalska-Kołaczyk 2002), namely a division into natural processes and learned rules. There, a set of diagnostics is set up to make the distinction, and since the learned rules start off, diachronically speaking, as natural processes, then any rule for which this historical path cannot be shown should be excluded by the model. According to Anderson (1985: 342), however, there is no formal way of making the distinction between rules and processes, and therefore between possible and impossible rules.

tongue root). The single-feature analysis has the advantage of obviating such stratagems and dispensing with the “ersatz vowel height feature” (Clements & Hume 1995: 283), i.e. $[\pm\text{ATR}]$. After evidence has been gathered for the articulatory reality of $[\text{ATR}]$ for West African languages, it has also been applied to vowels of Indo-European languages {Kenstowicz 1994 #548, or even consonants (sic!) in Slavic (Ćavar 2007). It seemed preferable to $[\pm\text{tense}]$, for which no articulatory evidence could be found. The relevance of the feature $[\pm\text{ATR}]$, however, for Indo-European languages is highly questionable (see Moosmüller 2007 for further discussion), and an analysis dispensing with it is to be valued. For a possible representation of vowel height of ME vowels employing only the feature $[\text{open}]$, see Figure 1. This representation seems to be well-suited to the four-way height contrast of the ME vowel system.

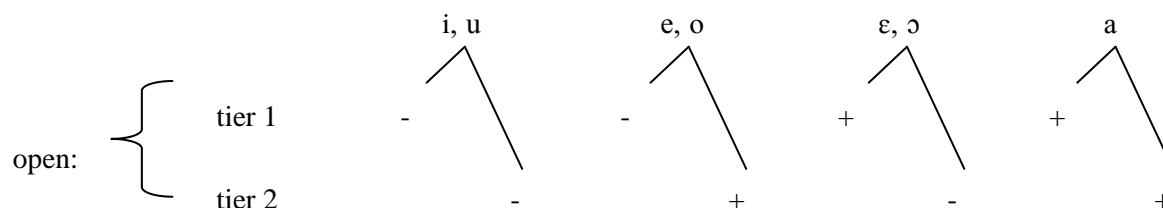


Figure 1: The representation of vowel height of ME vowels using the feature $[\text{open}]$

In addition to place features and height, the third feature relevant for the description of English vowels is that of vowel length. The difference in quantity between long and short OE and ME vowels and consonants can be represented by linking their root nodes to either one or two nodes on a separate tier representing quantity (see Figure 2). The question of what these nodes exactly represent, i.e. whether they should be seen as a reflection of the basic CV alternation (Clements & Keyser 1983), a universal alternation (Hyman 1982), or moraic structure (Hyman 1985) remains open, but that vowel length is best represented by linking the root node to two higher level entities of some sort is mostly uncontested (Odden 2011).

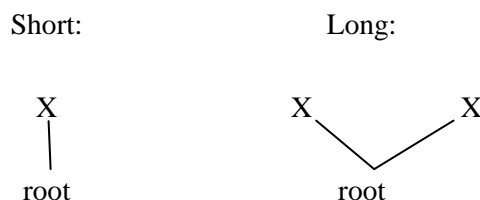


Figure 2: Representing vowel length (after: Clements & Hume 1995: 256)

While it has been argued above that the use of a single feature [open] is preferable when representing a four-level vowel length system of Middle English over the use of additional features, such as [\pm ATR] or [\pm tense], it is still necessary to introduce another feature, namely [\pm tense] to describe vowels of English after the Middle English period. The vowel pairs differentiated solely by length up to ME, have come to be additionally differentiated by quality after the exaptation of qualitative differences between the members of these vowel pairs. Thus, there clearly continue to be two classes of vowels in English, with some restrictions imposed on their distribution, and after the loss of reliability of duration as a correlate of length in maintaining the contrasts, length alone does not suffice to characterize this situation. Hence, starting with EModE onwards [\pm tense] is employed as a primary feature distinguishing the two classes of vowels. The following paragraphs present further argumentation for this choice.

Most accounts of the PDE vowel system agree that both length and tenseness¹⁵ are relevant features in its description¹⁶. They differ, however, as to which of these features is primary and which secondary. Consequently, there is no agreement in textbook accounts with regard to an important aspect in the representation of English vowels. Giegerich (Giegerich 1992) and Gussmann (2002), for instance, follow Chomsky and Halle's (1968 [1997]) use of the feature [\pm tense]. Harris (1994) and Hammond (1999) follow the view of later generative literature, such as Lass (1976) and Halle and Mohanan (1985), both referred to in Harris (1994), by giving primacy to length. McMahon (2002) maintains that either feature could be chosen, with no benefit of one choice over the other.

That some such feature is needed, and that it makes possible the discussion of the English vowel 'system' rather than 'inventory' is clear from a lucid presentation in Giegerich (1992: Chapters 4 and 6). He gives the following arguments for the existence of several 'pairs of phonemes'. There are a number of pairs of vowel phonemes in English which (a) are similar in quality (b) enter into long/short oppositions (c) remain in a 'semi-complementary' distribution, and two further arguments are that (d) there are several pairs linked by the same criteria, and (e) such pairs are collapsed in some varieties in a similar way (Giegerich 1992: 48). Giegerich (1992: 99) not only argues that a phonological feature should be postulated to

¹⁵ or ATR since Halle and Clements (1983) subsume [\pm tense] under [\pm ATR]

¹⁶ For a concise overview see: Carr & Honeybone (2007: 138).

distinguish these pairs of vowels, he also specifically argues that the feature in question be tenseness rather than length. First, to the objection that tenseness is ill-defined in phonetic terms he responds in kind by noting that length is also a questionable choice from the phonetic standpoint, since the ‘short’ vowels can have longer durations than their ‘long’ counterparts. This has repeatedly been shown before, cf. also Section 4.5.1. Second, he refers, without references, to perceptual studies which show that native speakers rely on qualitative differences more so than on duration when distinguishing between these vowel pairs. For a review of those, see: Bogacka (2003). His third argument for choosing tenseness rather than length is dialectal, namely that $[\pm\text{tense}]$ can be used to specify more varieties of English than $[\pm\text{long}]$, in view of the absence of the latter from Standard Scottish English, for example. While the last argument might in the end have more to do with the elegance of a cross-dialectal description than with legitimacy of choosing the right feature for the characterization of individual varieties of English, the first two arguments make a strong case that tenseness, rather than length, is a feature operative in the English vowel system.

To account for the fact that, in the same context, the vowel pairs are distinguished by duration, which points to the fact that length is redundantly predictable from tenseness, Giegerich (1992: 141) refers to syllable structure. Under the view he proposes, the tense vowels are associated to two x-bar positions on the timing tier, and the lax vowels are associated to only one x-bar position. This association takes place through late rules in a rule-based framework. This move, unavailable in classical OT, can be translated into a constraint holding at a higher level in Stratal OT, demanding that the specifications of segments for $[\pm\text{long}]$ and for $[\pm\text{tense}]$ have the same value.¹⁷

To conclude, it is thought of as a viable option to assume a single feature height specification by means of the single feature [open] for the four-way vowel contrast of ME. This is, however, not necessary for the later stages of English, where the height differences follow from the independently needed $[\pm\text{tense}]$ feature anyway. $[\pm\text{tense}]$ is regarded as a primary feature distinguishing between vowel pairs of similar quality in PDE, with length

¹⁷ Hammond (1999), in his cursory treatment of the representation of English vowels uses the fact that tense vowels are associated with two timing units as an argument in favor of length over tenseness. This is arguably the right conclusion if one assumes a strictly parallel view of OT, under which no equivalent of late rules can be introduced. Under Stratal OT, though, a segment with the feature [long] in its featural specification is not necessarily more easily attachable to two timing units than a segment with the feature [tense].

differences being secondary. As for the front/back dimension, a traditional single-value specification with [\pm back] suffices to specify English vowels for all stages presented here.¹⁸

2.3.2.2 Preservation of contrast

After a slight detour in the questions of segmental representations the discussion now moves back to a strictly OT concern, namely that of the preservation of contrasts. This concern is firmly rooted within OT, as one of the two main types of constraints, the faithfulness constraints, are responsible for maintaining lexical contrasts, and so this task is built into the very core of grammar in OT.

In the analysis proposed in the present thesis, the issue of contrast is of paramount importance, since the preservation of lexical contrasts is seen as a force in the avoidance of mergers. In classical OT (Prince & Smolensky 1993 [2002]), contrasts fall out of an interplay between markedness constraints on sounds and sound sequences and faithfulness constraints. It is the task of the faithfulness constraints to preserve contrasts, and whenever faithfulness outranks markedness with respect to a given contrast present in the input, it is preserved in the output.

An alternative solution is proposed by Flemming's (2004) Dispersion Theory of Contrast. He convincingly argues that it is not only individual sounds or sound sequences that are more or less marked, but that also contrasts between sounds can be evaluated in this regard. One of the examples that he analyzes is that of high central vowels. A classical OT solution to accounting for a lack of [ɨ] in English would be to set up a markedness constraint against it, such as *[-back, +high, -low, -round]. However, this vowel is not, according to Flemming, in itself marked, as evidenced by its surfacing when front and back is not a distinctive feature. This is for instance the case with the unstressed, word internal vowel in American English, which does not contrast with any other vowel in this position (Flemming & Johnson 2007). In fact, in such cases, a high central vowel seems to be the default, unmarked option. The vowel is not in itself marked, then, but it can become a member of a marked contrast. For example, in a phonemic inventory which contains other high vowels, namely [i] and [u], the

¹⁸ With schwa, that is the vowel appearing in unstressed position, lacking specification for high/low or back/front.

markedness of [ɨ] consists in the fact that it happens to provide a less salient, and therefore more marked, contrast to the other high vowels, then these do to each other. In other words, both contrasts involving high central vowels, that is a contrast between [i] and [ɨ], and a contrast between [ɨ] and [u], are marked in comparison to the contrast between [i] and [u].

To incorporate the role of contrast in shaping segmental inventories into the grammar, Flemming (2004) introduces a new family of constraints, adding to the two types of constraints (markedness and faithfulness) postulated in Prince and Smolensky (1993 [2002]). Members of this new family are constraints on the distinctiveness on contrasts, namely MINDIST and MAXIMIZE CONTRASTS. The requirements of those two types are to a certain extent contrasting. While there is only one MAXIMIZE CONTRASTS constraint, which demands that the number of contrastive sounds be as high as possible, there is a whole host of MINDIST constraints. They demand that distances between neighboring phonemes (in a given acoustic dimension, say F1 or F2) be not lower than a certain value. This acoustic distance is not measured in absolute values but derives from distances between sounds in the “coarsely quantized three-dimensional vowel space” (Flemming 2004: 10), which allows a speaker-normalized scale which abstracts away from the actual phonetics of individual realizations. The interaction between the MAXIMIZE CONTRASTS constraint with the various MINDIST constraints and, additionally, constraints on the markedness of individual sounds are shown to be crucial to neutralization (on the example of vowel reduction in Italian and English Flemming 2004: 16) and blocking in harmony processes (on the example of the Johore dialect of Malay Flemming 2004: 46). Although this particular formulation of the Dispersion Theory of Contrast may have its shortcomings, it is recognized that its pursuit to capture the relative markedness of contrasts is a legitimate undertaking and that the introduction of constraints on contrasts to do so is likely to be the right way of going about it (Steriade 2007); it will be kept for now, in the form of the PC theory proposed by Łubowicz (2003; 2011), which is presented below.

The finding that there can be constraints on contrasts is picked up by Łubowicz’s (2003; 2011) Preserve Contrast (PC) theory. She suggests the incorporation into the OT architecture of Preserve Contrast constraints, which she defines as follows (where P stands for “a

potentially contrastive phonological property, such as a distinctive feature, length, stress, presence vs. absence of a segment” Łubowicz 2003: 18):

(2) PRESERVECONTRAST: PC(P) (Łubowicz 2003: 18)

For each pair of inputs contrasting in P that map onto the same output in a scenario, assign a violation mark. Formally, assign one mark for every pair of inputs, in_a and in_b , if in_a has P and in_b lacks P, $in_a \rightarrow out_k$, and $in_b \rightarrow out_k$.

“If inputs are distinct in P, they need to remain distinct.”

PC constraints are somewhat similar to, and yet distinct from faithfulness constraints. Like faithfulness constraints, they consider the output with respect to its similarity to the input. Yet, PC constraints are distinct from faithfulness constraints in that they are satisfied by pairs of output forms even if they are different from the input forms (unlike faithfulness constraints) as long as contrast between the words is maintained. They take into account correspondences between entire words, and not between individual mappings and they administer violation marks to output words in which contrasts are neutralized, but they do not penalize the output if the contrast is realized differently in the output than in the input.

These constraints can also be employed to account for the preservation of contrasts in historical change. Successive generations of speakers construct PC constraints for which there is evidence in ambient language. They might, however, maintain the contrast by different means than the previous generation, if the previous means of maintaining it prove unreliable, as was the case with duration for maintaining the length contrasts in English. Namely, it can be argued that when language learners were faced with ambient language in which the old quantity-based oppositions /e:/ - /e/ and /o:/ - /o/ were maintained by a rather unreliable duration, the construction of PC constraints played a role in the preservation of these contrasts, albeit not by means of length anymore. Raisings of /e:/ and /o:/, as well as lowerings of /o/ and /e/ did not violate PC(length) constraints (although these two processes resulted in violations of faithfulness constraints). In other words, the relevant PC constraints were constructed by learners but since the contrasts were not robust enough, they came to be realized by different means.

2.4 EVALUATION OF APPROACHES TO LANGUAGE CHANGE (AND VOWEL SHIFTING) IN THE HISTORY OF ENGLISH PHONOLOGY

2.4.1 *Introduction*

The Great Vowel Shift is one of the most widely studied events in the study of English historical phonology. Ever since its formulation by Luick (1921-1940) and Jespersen (1909–49) it has attracted a great deal of attention, not least because it poses some fundamental questions about the nature of language change. One such question concerns the triggering of the change. As in any change, the actuation problem has not been definitely solved here. The question of why it took place at a particular time in the history and not another has not been answered, and the validity of asking such a question has been called into doubt. Another important question is this: can such a comprehensive reorganization of the sound system be conceived of as a unitary event? This question has sparked the debate between Stockwell and Minkova denying it on the one hand, and Lass defending it on the other. Both points are relevant for the following discussion.

The actuation problem is tackled in the account proposed in this thesis by means of placing the shift in the context of an even more global reorganization of the rhythmic system of English, an increase in the degree to which English has a stress-based rhythm. The question of the coherence of the shift is defended from the standpoint of an evolutionary perspective, where environment can exert selection pressures over generations of speakers. Such a conceptualization overcomes the problem inherent in functional explanations invoking the needs of speakers.

The following sections sketch the different approaches to language change in general and to vowel shifting in particular. First, Section 2.4.2 sketches the characteristics of the Neogrammarian approach to language and language change, introducing the issues of the regularity of sound change, the role of speakers in directing the change and the role of language acquisition vs. language use for language change. Section 2.4.3 discusses structuralist/functionalist accounts, which introduce the recurring question of the needs of speakers as a factor in language change, including the need to preserve lexical contrasts, and the need to minimize the articulatory effort. Section 2.4.4 presents how the Great Vowel

Shift has been dealt with within the classic generative work, SPE. Section 2.4.5 presents how vowel shifting has been approached within OT. Section 2.4.6 introduces exemplar-based phonology, whose insights are seen as very useful in tackling the question of the effects of language use on language structure and language change.

2.4.2 The Neogrammarians

The Neogrammarians, as a reaction to the rather metaphysical bent of the philosophy of language of the time, strived in their pursuits after scientific objectivity. They did that by clearly delineating the object of study in linguistics by tying it to observable phenomena. This is laid out in Hermann Paul's (1880 [1920]) *Prinzipien der Sprachgeschichte*, considered the 'handbook' of the Neogrammarians (Kiparsky 1965; Murray 2010). Following from this conviction is the limiting of investigations to the idiolect, or actually the collection of all idiolects of all speakers, rather than the underlying, linguistic system of a social nature. According to Paul (1880 [1920]: 24), earlier, purely descriptive grammars tried to describe particular stages in the histories of those languages, and sought to present their histories by producing a sequence of such grammars. What is problematic is that such grammars of particular stages were merely abstractions over all the speakers of respective speech communities. Since abstractions cannot interact, such a treatment of the history of a language is a misrepresentation. The entities which undergo change are "die psychischen Organismen", the mental representations of linguistic forms, since the actually occurring utterances cannot change (Paul 1880 [1920]: 28). A spoken utterance vanishes as soon as it is pronounced, and a written utterance remains forever unchanged, so neither can be said to have a history. There is no direct causal relationship between successive pronunciations of what is thought of as the same sound; there is a relationship between the mental representation and each individual pronunciation, and in the longer term, between the mental representation of one speaker and that of another. A further consequence of focusing on the observable phenomena for the linguistic practice is that most of research conducted by the Neogrammarians was limited to the most directly accessible level of language, namely its sounds.

The third chapter of Hermann Paul's (1880 [1920]: 49) is a good outline of the Neogrammarian point of view on the sounds of language and on sound change. It presents a

mechanism of sound change which is rooted in physiology, and is therefore thought to overcome the problematic status of abstractions interacting with each other. The Neogrammarians worked on the assumption that speakers are not consciously aware of the sounds they produce. Even though a certain level of consciousness can be stipulated for the process of acquiring sounds, being it first language acquisition or second language acquisition, a fluent speaker of a language is no longer fully conscious of the language in their command, and in fact, this lack of conscious attention, or its transition from declarative to procedural memory as it would be stated in modern parlance (Anderson 1993, Boyland 1996, referred to in Bybee 2001: 8), is what makes fluent use of language possible. As a consequence, change was also seen as happening below the level of consciousness of speakers. Its locus was seen in the unintentional variation in speech production, which influenced the mental representation of a sound. Paul uses the term *Erinnerungsbild* ‘memory trace’, as a label for the entirety of the representation of a sound, which consists of a *Bewegungsgefühl*, i.e. an articulatory representation, and of a *Lautbild*, i.e. an auditory representation. In a manner presaging modern exemplar theories (cf. Section 2.4.6), Paul characterizes the storage, and change, of speech sounds in the following way:

dies Gefühl [ist] das Produkt aus sämtlichen früheren bei Ausführung der betreffenden Bewegung empfangenen Eindrücken, und zwar verschmelzen nach allgemeinem Gesetze nicht nur die völlig identischen, sondern auch die unmerklich von einander verschiedenen Eindrücke mit einander. Ihrer Verschiedenheit entsprechend muss sich auch das Bewegungsgefühl mit jedem neuen Eindruck etwas umgestalten, wenn auch noch so unbedeutend. Es ist dabei noch von Wichtigkeit, dass immer die späteren Eindrücke stärker nachwirken als die früheren. Man kann daher das Bewegungsgefühl nicht etwa dem Durchschnitt aller während des ganzen Lebens empfangenen Eindrücke gleichsetzen, sondern die an Zahl geringeren können die häufigeren durch ihre Frische an Gewicht übertreffen (Paul 1880 [1920]: 55).

In short, a speaker carries a representation of a sound in their mind. The articulatory representation is influenced by the actual instances of pronouncing the sound. This auditory representation can change, when the variation inherent in the production of the sound drifts in a particular direction. The change in the articulatory representation then causes a change in the auditory representation.

What is also claimed in the above-given quotation is that not all instances bear the same weight in influencing the characteristics of the representation, and more recent

pronunciations contribute more to the end product. This is precisely what enables a new variant, stemming from the inherent variation in speech production, to gain a foothold and come to influence the representation of the sound. The variants that tend to prevail are those that result in greater *Bequemlichkeit*, which can be understood as greater ease of articulation. For sequences of sounds, assimilation is an obvious example, where sequences of more similar sounds are easier to pronounce than sequences of more different sounds. For individual sounds, however, it is not possible to say which pronunciation is easier, and the ease is dependent on other sounds in the system of a particular language. Small modifications of sounds can thus take place within the lifetimes of individuals, but more substantial changes happen when a language is passed on from one generation to the next.

At least two points of comparison with modern exemplar theories are worth mentioning here. First, the higher weighting of later forms is not normally modeled by exemplar approaches to phonology, but in principle it could well be incorporated. Second, another, crucial claim by Paul, namely that the representation of a sound which is common to all its instantiations in all words changes as a whole, not only of a sound in particular words, signals an important problem to exemplar-based phonology. Changes of the sort postulated by Paul, that is changes in which a given sound changes in all words in which it appears, are attested. Whereas exemplar theory is well-suited to account for the cases of word specific change of very high or very low frequency items, it is challenged by the regularity observable at least in some sound changes.

The issue of regularity in sound change is, in fact, a controversy started by the Neogrammarians, which in some form or another continues to the present day. Under the Neogrammarian conception, sound change is completely regular.¹⁹ This very categorically expressed idea has probably remained one of the most widely discussed Neogrammarian ideas, at least with regard to phonology. Schuchardt (1885: 34) traces the idea of exceptionlessness of sound change to August Schleicher, but one of its clearest articulations can be found in Brugmann (1885, cited in 1885: 23), who postulates that “bei dem Vollzug des Lautwandels ist nun gar nicht denkbar dass in verschiedenen Wörtern verschiedene Wege

¹⁹ Thus, for example, Verner’s Law, which restored the exception-free view of Grimm’s Law, was a welcome finding.

eingeschlagen werden”. Paul’s (1880 [1920], cited in Schuchardt 1885: 23) arguments for this position, are the following:

Das Bewegungsgefühl bildet sich ja nicht für jedes einzelne Wort besonders, sondern überall wo in der Rede die gleichen Elemente wiederkehren, wird ihre Erzeugung auch durch das gleiche Bewegungsgefühl geregelt. Verschiebt sich daher das Bewegungsgefühl durch das Aussprechen eines Elementes in irgend einem Worte, so ist diese Verschiebung auch massgebend für das namliche Element in einem anderen Worte.

This view of sound change as completely regular was criticized by the contemporaries, notably by Schuchardt, who observes that the frequency of occurrence of a word influences the likelihood with which it will participate in a sound change (more on the implications of this observation for phonology in Section 2.4.6.). Despite occasional voices of dissent, though, the view of sound change as regular has prevailed as the mainstream view, not only in the times of the dominance of the Neogrammarians, but also in the following years of structuralist linguistics (up until arguments for a mechanism resulting in irregular change, namely ‘lexical diffusion’ 1975 started cropping up, see next paragraph). According to Anderson (1985: 257), one of its most vehement proponents among the structuralists was Bloomfield, who famously rephrased the argument of the regularity of sound change from the Neogrammarian dictum ‘sound laws have no exceptions’ to a structurally-grounded one: ‘phonemes change’ (Bloomfield 1933: 351).

The question of whether phonemes or words are the units undergoing sound change has been later dubbed ‘the Neogrammarian controversy’ (Labov 1981). In Wang and Cheng’s (1975) terminology, the Neogrammarians postulated that sound change is phonetically gradual and lexically abrupt. They, on the other hand, claim that sound change is phonetically abrupt, but lexically gradual, thus postulating a mechanism of sound change which they called ‘lexical diffusion’. While lexical diffusion has since been recognized as a mechanism of sound change, it is not the case that it is the only possible way in which sound change can proceed. Labov (1981), for instance, states that it might complement, rather than replace, the Neogrammarian view. Under his conciliatory approach to the Neogrammarian controversy, the Neogrammarians were not really making a statement about what kind of change is possible in general, but they rather defined the kinds of change that they found worth investigating. And so, on this sympathetic reading of the Neogrammarians both types

of change are compatible even from their point of view, and it is just a matter of figuring out under what conditions each of the changes is to be expected. And so Labov (1981: 296), after a detailed analysis of two kinds of changes in Philadelphia concludes that the short-*a* split is an example of lexical diffusion, and that the vowel shifts are classic examples of Neogrammarian exception-free sound change. Trying to generalize the findings on sound changes in progress to historical changes, he sees qualitative changes such as vowel shifts, including the Great Vowel Shift, as consisting primarily of regular changes, and quantitative changes, i.e. shortenings and lengthenings, as cases of lexical diffusion. The proposed explanation for why some changes take the form of lexical diffusion, i.e. are irregular, is that these involve a change in the membership to an abstract category, such as tense or long, rather than just the modification of phonetic characteristics of a sound. Features such as tense are more abstract than those such as front, because they do not refer to a single physical dimension. In other words, changes between larger subsets of vowels are more likely to be irregular than changes within those subsets. Labov (1994 [2010]: 542) later identified regular change with ‘change from below’ and lexical diffusion with ‘change from above’, which suggests that a change might start off as regular, but end up as lexically diffused once it reaches the awareness of a speech community. A new perspective on the issue of regularity of sound change is forced by Bybee’s exemplar-based theory. Following from her assumptions that each lexical entry is linked to its own exemplar cloud and so each word has its own gradient phonetic specification, she concludes that the prevalent instances of sound change are both phonetically and lexically gradual (Bybee 2001: 40). The problems with this point of view are discussed in Section 2.4.6.

Paul’s theory is also often invoked with reference to another important issue in the mechanisms of linguistic change, namely with the role of language acquisition by children, versus the role of language use by adults. Though he is often cited (cf. Luraghi 2010) as advocating a first language acquisition based approach, his position on the question of whether sound change happens mostly through first language acquisition or through language use, Paul’s actual stance is somewhat equivocal. He admits of both possibilities, without clearly stating which one is to be conceived of as the principal one. On the one hand, he states that “[e]s liegt auf der Hand, dass die Vorgänge bei der Spracherlernung von der allerhöchsten Wichtigkeit für die Erklärung der Veränderung des Sprachusus sind” (Paul

1880 [1920]: 34). On the other hand, since change is argued for to be driven by articulatory considerations and the accumulated effect of articulatory modifications on the articulatory, and then auditory representations, ‘the articulatory drift’, then clearly change has to be postulated to take place within single generations of speakers, and it is not a trivial number of changes, but rather the principal one (Kiparsky 1965). This mechanism, one of entrenchment of the effects of language use on mental representations, is employed the proposal of this thesis, in that the frequent associations of vowels with strong prosodic positions are argued to have contributed to their raisings and diphthongizations, and frequent associations of vowels with weak prosodic positions are argued to have contributed to their reduction and loss. Further research has tended to side either with the child-based or adult-based view. Theories of Lightfoot (2003) or Stampe (1979), for instance, place the locus of change in children. Others, such as Drachman (1978), Hock (1992, referred to in Luraghi 2010), or Aitchison (2003), pointing to the dissimilarities between the kinds of substitutions found in child language and those that can result in historical change, question the validity of the child-based perspective. Additionally, some of the research in sociolinguistics which sees transmission, rather than innovation as the crucial component of linguistic change (e.g. Milroy and Milroy 1985), sees the role of language acquisition as negligible, if not nonexistent, since, as Aitchison (2001: 209) puts it, “[b]abies do not form influential social groups”. Despite this criticism, though, it seems premature to reject the role of language acquisition for language change. Even though children do adjust their speech patterns to those of their caregivers, peers, and speakers of whatever social networks (or communities of practice) they belong to, it is still conceivable that some features remain unchanged. This is especially plausible with changes to the grammar which are hardly visible in surface forms, such as reanalysis or rule reversal, and which are therefore unlikely to attract social opprobrium, which would lead to their suppression as the babies involved interact with others. Also, the later adjustments might remain only formal speech variants, with the innovative forms surfacing in casual speech settings. And if that happens, then the speaker, no longer being a baby, and so possibly part of an ‘influential social group’, might become a vector of language change.

To conclude, the Neogrammarians staked out some important claims which remain important in the discussion of language change to this day, and their take on those is in many

cases adopted in the present discussion. In particular, Paul's take on the entities undergoing sound change as lines of successive mental representations of sounds is at the very core of the evolutionary perspective adopted in the present thesis, which sees lineages of linguistic replicators as changing. More generally, the importance of diachrony for explaining language structure is also shared. Another staple of Paul's thinking, namely the assumption that speakers as agents cannot be brought to bear in explanations of language change is also an important component of a replicator-based account. Additionally, the importance of physiological factors for the direction of sound change is assumed as well. As for the role of children versus adults in language change, in line with Paul's position, both possibilities are admitted, with the entrenchment of frequently recurring speech patterns exerting influence on the kind of input that forms the basis for the language acquisition of future generations.

2.4.3 The structuralist/functionalist approach

The seminal work of de Saussure (1916 [1995]) started the movement in linguistics that has come to be known as structuralism. Various scholars who came after de Saussure took up a number of dichotomies established by him; the ones having the most direct relevance for the study of language change are those of synchrony vs. diachrony, and of *langue* vs. *parole*. Saussure advocated a change in focus in the study of language from diachrony to synchrony, taking as his starting point the observation that language is a system of oppositions, forming a coherent whole at any stage in its development. This stands in contrast to the investigation of individual sounds changing through time, without regard to their role in the linguistic system, which is how he saw the practice of Neogrammarians, a practice deemed as 'atomistic'. Thus, for de Saussure, it is the synchronic stages of languages that should be studied and not diachronic developments, since he saw them as investigating the relationships between arbitrary pairs of linguistic systems, and therefore uninteresting with regard to the description of those systems themselves

The radical shift from diachrony to synchrony was somewhat weakened already in the work of the Kazan school (especially of Jan Nieciśław Baudouin de Courtenay), and the Prague school (notably Roman Jakobson), where the links between synchrony and diachrony were re-instated. According to de Courtenay, for instance, the phonological alternations which do not have a synchronic motivation at a given stage in the development of a language

(in his terminology: ‘correlations’) do come into being as synchronically motivated at an earlier stage (in his terminology: as ‘divergences’) (Baudouin de Courtenay 1895). Another link between synchrony and diachrony was postulated by Jakobson (1944 [1969]: 79), according to whom the universal implicational laws which are at play in language acquisition also limit the possible structures present in languages, and consequently restrict the direction of possible sound changes.

Apart from linking diachrony and synchrony, another contribution of Prague school linguists to investigating language change is their focus on function, next to structure, as an important aspect in linguistic theorizing. For them, particular elements of the linguistic system, such as phonemes or morphemes, were defined not only by their position in the structure of the system by standing in opposition to other elements, but also by the function they fulfill. The tradition of linguistic thinking which shares this focus, namely functionalism, is marked by the conviction that language structure reflects language use, that is that ‘form follows function’ rather than ‘function follows form’, as is articulated by the programmatic pronouncements of Mathesius:

The new linguistics conceives language as something living, underneath the words it sees the speaker or the writer from whose communicative intention they have resulted (Mathesius 1983: 122)

the new linguistics (...) starts from the needs of expression and inquires what means serve to satisfy these communication needs (...) It thus proceeds from function to form (Mathesius 1983: 123)

Mathesius even saw the change of focus from phonetics, to phonology as part of the functional turn, since he proclaims: “whereas phonetics studies sounds, phonology studies phonemes, i.e. sounds endowed with functional meaning” (Mathesius 1983: 129). The advent of synchronic phonology is tied to another linguist associated with the Prague Circle, namely with Trubetzkoy, and the functional aspect of phonological analysis is particularly clear in the work of André Martinet, who was very much inspired by Trubetzkoy’s approach to phonology.

This focus on function, next to structure is elaborated on with regard to sound change most famously by Martinet (1952), for whom both the structure of languages and the processes of change that they may undergo are a reflection of the needs of their speakers. In

his account of vowel chain shifting, Martinet proposes a mechanism of causally related changes, where the causality lies in the need of the speakers to avoid merger, once one vowel encroaches on another. Namely, speakers do not want to merge word classes, and so, to avoid homonymy, they come to change the quality of another vowel as well.

The way of approaching language change as proposed by Martinet (1952) incorporates functional, structural and physiological factors in the process of a vowel chain shift. In terms of functional factors, the communicative needs of the speakers are evoked. These needs (e.g. being understood or drawing attention to a specific portion of the message) have to be satisfied for communication to work, and they tend to maximize contrasts. Communicative needs stand in opposition to physiological factors (principle of least effort, range of dispersion), which tend to minimize contrasts. Additionally, on the structural level, there are the pressures of the linguistic system (e.g. pattern attraction) that influence the direction of the change. The interplay of those three sets of factors results in change. For chain shifts specifically, Martinet adopts an account which is in a way similar to that of Luick's 'displacement theory', who describes his position in the following way:

Wenn [...] in der Umgrenzung des Phonems ein psychologischer Faktor wirksam wird, so ist es nicht überraschend dass auch ein Gefühl für den Abstand zwischen den Phonemen besteht und ferner eine Tendenz, diesen Abstand zu erhalten. [...] Wenn also ein Phonem in Bewegung gerät, wird häufig auch ein anderes davon berührt, und so kann es zu einem "Verdrängen" kommen (Luick 1932: 89, quoted in Stockwell & Minkova 1988b: 357).

Martinet's proposal also involves causality between the movement of one phoneme and that of another, but he puts forward an explicit mechanism, which involves 'the communicative needs of the speakers', which is supposed to account for it. He suggests that there is inherent range of different phonetic realization of any given phoneme (even in the same environment), referred to as 'range of dispersion', with the most frequent realizations forming 'the center of gravity'. Between the ranges of contiguous phonemes lie 'margins of security'. When one phoneme moves, the margin of security between it and an incoming phoneme increases, the range of dispersion of the newcomer changes, and consequently its center of gravity also shifts. The existence of the range of dispersion is clearly a physiological fact, grounded in the limits on the speech organs, the move of the center of gravity, however, is already a systemic change. Additionally, to account for the questions of why mergers do not occur in such a

scenario, the issue of ‘functional yield’ is brought in. Speakers avoid mergers if these would destroy a contrast which is important for maintaining distinct a large number of lexical items. This is a mixture of a systemic feature (the number of lexemes kept distinct by a given phonological contrast) and the needs of speakers to maintain this contrast. The prediction for when a merger, rather than a chain shift, takes place is that a contrast with a low functional yield is more likely to fall prey to a merger than a contrast with a high functional yield.

This answer to the question of why, in a particular instance of language change, a shift, rather than a merger takes place is not definitive, however. Martinet’s solution to the problem of determining when a merger rather than a chain shift is to be expected by means of invoking functional yield has been severely criticized. King (King 1967), for example, having derived testable hypotheses by turning the notion into numeric formulas, shows that functional yield (or ‘load’) is not a good predictor of the developments in three languages, namely Icelandic, German and Yiddish. Also, it has a hard time dealing with mergers which abolish very functional contrasts, such as the *meet/meat* merger in English. Still, being able to express semantic contrasts, which is an important function of any natural language, hinges on the presence in those languages of a good deal of formal contrasts. Consequently, even if functional load does not make correct predictions with regard to language change, the idea that it is a property of grammar, and a historically observable tendency, to retain contrasts, and that it might be more important to maintain a contrast the more ‘useful’ it is, has not vanished from linguistics. For instance, the need to preserve contrasts is embedded into the architecture of OT as embodied in the presence of faithfulness constraints (and of the PC family of constraints proposed by Łubowicz 2003, see Section 2.4.5). The question of why a chain shift rather than a merger happens is very much relevant to the present investigation, since it tries to model a mechanism where a set of vowel shifts was put in motion. The question of why a set of shifts, and not a set of mergers, took place is an important one. To address it, Appendix 1 presents a comparison of developments in Romance and Icelandic, where major reorganizations of vocalic systems took place, resulting in both mergers and vowel shifts. It suggests that when a phonological feature disappears from a language, then qualitative shifts, rather than wholesale mergers, tend to take place. With a large number of vocalic oppositions, this is likely to result in chain shifts.

Next to the communicative needs of the speakers, which find their expression in the maintenance of contrast and keeping linguistic forms distinct, functional approaches to language recognize the existence of another point of contact between language form and language use, which stems from the fact that speakers also strive to minimize the effort expended when speaking. This striving is the need to minimize articulatory effort. It has been common at least since Passy (1890) to maintain that phonetic change leads to greater ease of pronunciation, thus reflecting the principle of least effort, namely the principle that human behavior in general, including language specifically, is guided by the tendency to minimize effort (which has found one of its most well-known formulations in Zipf's 1949 monograph).

Given that an account of linguistic change can refer to these two opposing kinds of needs, the need to communicate, and so to maintain contrasts on the one hand, and the need to minimize effort, which involves losing contrasts on the other hand, can prove too powerful to be explanatory in the end. For instance, Samuels (1972, cited in Stockwell & Minkova 1988b: 357), with reference to the developments in the English vowel system, suggests that the diphthongization of high long vowels, which under his account started the GVS, is an effect of "forceful style variants", whereas the laxing of short vowels started in "relaxed style variants". Similarly, Leith puts the raising of the *mate* vowel down to 'forceful speech variants', and the diphthongization of the high vowels to "a 'lax' pronunciation of the diphthongs as a means of economizing on articulatory energy" (Leith 2002: 143). As such, almost anything can be explained by recourse to the contrastive notions as 'ease of production', 'ease of perception'. This adds to the already mentioned problematic issues of functional explanations which stem from their assumptions about human agentivity. An additional problem, namely the lack of clarity about what is articulatorily difficult and what is easy is made painfully clear by the fact the Samuels and Leith 'explain' the same development, namely diphthongization of high vowels by recourse to increased effort and decreased effort, respectively.

Precisely this tension between two opposing forces active during language use, which then influence language change lies at the heart of two phonological frameworks, which are both functional, namely Natural Phonology and Optimality Theory. In Natural Phonology, where language is thought to be "a natural reflection of the needs, capacities, and world of its users" (Donegan & Stampe 1979), the need to maintain contrasts manifests itself in the form

of foregrounding processes, and the need to minimize effort in the form of backgrounding processes. The resolution of the conflict boils down to a particular ordering of the application of the respective processes. In OT, on the other hand, the need to maintain contrasts translates into faithfulness constraints and the need to minimize effort into markedness constraints. The resolution of this conflict results in a particular ranking of the respective constraints.

Even though the functionalist approach introduces important issues, it is far from satisfactory. The question of when it is legitimate to invoke the needs of the speakers, for one, remains largely unanswered. While both the physiological aspects and the systemic aspects can be investigated, it seems unlikely that limits as to what speakers may want can be circumscribed. In any case, the needs of the speakers seem problematic in an approach that makes them so prominent, since the cases where changes abolishing useful contrasts happen clearly contradict the all-powerful status of speakers in shaping their language. Seeing speakers' needs as only one of many selective forces on the evolving system, on the other hand, makes clear why sub-optimal changes are free to happen (Ritt 1995). This change of perspective, slight though it may seem, is a crucial step in devising a coherent view of language change, where there are no almighty speakers re-arranging their language, but where speakers are seen as hosts of languages and their needs might influence the process of replication, but not orchestrate it. A case in point is the stipulation, admitted by Martinet (1952: 10), that a low functional load does not necessarily lead to the loss of contrast, with /ʃ/ and /ʒ/ remaining distinct in English, despite the nonexistence of contrasts kept distinct by these two sounds. If the tendency to minimize effort were a driver of linguistic structure, kept in check only by the communicative function embodied in the functional load, though, then the maintenance of contrasts with low functional loads would have to be seen as perverse.

Summing up, the functional considerations proposed by the scholars mentioned in this section, an example of which is the need to preserve contrasts formulated as the hypothesis of functional load by Martinet, or the need to minimize effort, do play a role in language change. As such, it is a welcome extension of the factors discussed by the Neogrammarians. However, their incorporation into linguistics is problematic if framed in a speaker-based perspective. At any rate, the 'functional' nature of language change, that is the stipulation

that there are instances of language change which improve languages in their efficiency as tools for communication is taken to be a possibility, but by no means the only or even the main driving force in language change. For each such instance, it would have to be shown how and why the greater efficiency of the new form has come to outweigh whatever advantages the old form had that made it dominant, for the time that it was dominant. Functionality is only one of the environmental factors to be considered in an evolutionary account.

2.4.4 *Generative phonology*

In order to discuss the treatment of language change in generative phonology, it is important to first go back to the already mentioned distinctions set up by de Saussure. One of the main contributions of de Saussure, to linguistics in general and to phonology specifically, is the introduction of the distinction between ‘langue’ and ‘parole’. ‘Langue’ refers to the shared conventions of the speech community, and as such is social in nature. ‘Parole’, on the other hand, refers to individual instantiations of this shared system in actual spoken communication, and as such is individual in nature. For de Saussure, it is the generalities of ‘langue’ rather than the particularities of ‘parole’ that should constitute the object of study in linguistics. According to him, phonological change stems from facts of parole, and is then taken up and spread by a few individuals, before the entire speech community might take it over. As such, it is consigned to parole and banished from linguistics (Anderson 1985: Chapter 3).

Generative linguistics takes up de Saussure’s dichotomy of ‘langue’ and ‘parole’ and sees the implications of this dichotomy to be as far-reaching as de Saussure himself did, albeit in a slightly different way. The dichotomy corresponds to Chomsky’s (1965) ‘competence’ vs. ‘performance’, and later (Chomsky 1986) ‘I-language’ vs. ‘E-language’ distinction. The strict separation of the linguistic system and language use suggests that describing language change amounts to finding links between various linguistic systems, which happen to constitute various stages in the history of a language. The task of linguistics is to describe those individual stages; the fact that they are related by inheritance is, at most, of secondary importance. However, in contrast to de Saussure’s ‘langue’, competence (or I-language) is

not seen a social fact, but rather as a property of individual, even if idealized, speakers/hearers (Anderson 1985).

One of the first treatments of phonological change in generative grammar is Halle's (1962) paper, wherein he proposes that "the primary mechanism of phonological change is the addition of rules to the grammar with special (though not exclusive) preference for the addition of single rules". Thus, he explicitly states what is a logical conclusion from Chomsky's idea that languages are grammars instantiated in minds of their speakers. If languages are grammars, then linguistic change must be a change in grammars. Since for classic Generative Phonology, as for many further conceptualizations, grammars consist essentially of rules, then a change in grammar amounts to a change in the rule system. For Halle, the simplest of the possible changes is the addition of a new rule.

The treatment of sound change generally, and of the Great Vowel Shift specifically, expressed in the *SPE* following from this premise is a rather idiosyncratic one. Interestingly, the account of English phonology provided therein claims that the changes in the phonology of English that took place during the GVS did not alter the underlying phonemic representation of words, and it is in fact one of the goals of Chomsky and Halle's treatment of the evolution of the English vowel system to "provide some explanation for the remarkable stability of the underlying system of representations" (Chomsky & Halle 1968 [1997]: 249). The explanation of the sound change which has resulted in the current shape of English vowels is provided as a case of rule addition, the most straightforward case of describing differences between different stages of the same language. When a rule is added to the phonology of a language, the representations might well remain unchanged. The representations are claimed to be the same for living speakers as they were for pre-GVS English. What accounts for the difference in the surface forms of pre-GVS and post-GVS English is an addition of a rule, the Vowel Shift Rule (Chomsky & Halle 1968 [1997]: 187), to English phonology. As a result, current surface forms are arrived at by adding one more rule in the derivation performed on unchanged underlying forms.

The conceptualization of rule addition as the main mechanism of language change might suggest that languages accumulate more and more rules as they grow older, with all the younger stages of the languages present in the rule systems of their present-day descendents.

However, it should be noted here that languages do not have to be living repositories of all rules ever added to them under the generative view. What is supposed to act against such accumulation are cases in which children construct grammars which produce the same surface forms as the grammars of their parents, but which lack a particular rule of the older system, since “[t]he children’s grammar will contain a given phonological rule which corresponds to a historically attested change and is present in the grammar of their parents only if the grammar containing this rule is the most highly valued grammar in terms of the evaluation measure” (Chomsky & Halle 1968 [1997]: 251). Since grammars are evaluated according to their simplicity in generative linguistics, this simply means that children can construct simple grammars yielding the same output, in which case a rule is lost from the language.

With reference to the question of the coherence of the chain shift, the *SPE* provides an account under which the individual qualitative changes are undeniably related. It is the same rule which applies to all members of the shift which is responsible for the divergence of their surface forms from underlying representations, and so the Great Vowel Shift clearly has unity and coherence under this view. As for what kind of event it was, it is conceived of here as the addition of the vowel shift rule in the competences of speakers of English. Aside from locating the historical change within the synchronic phonology of speakers, this approach does not attempt to provide reasons for the addition of the rule, and so it is, in that sense, purely descriptive. The Vowel Shift Rule was added to the phonology of English, and questions as to why this happened are outside the range of issues tackled by the theory.

To go back to the mechanisms of change under the generative paradigm and to the relation between synchrony and diachrony therein, the unification of synchronic and diachronic investigations within the generative paradigm was taken up by Paul Kiparsky, a staunch believer in the relevance of diachronic investigations for synchronic phonology, who aptly formulates the goal of generative research on diachronic phonology as follows.

The strategy is to identify structure-dependent properties of change and to use them in turn to test hypotheses about structure. For example, if the right way to look at analogical change is not as the projection of surface regularities but as the elimination of arbitrary complexity from the system, in a sense of complexity independently defined in the theory of grammar, then it follows that particular instances of change

can show something about the grammars of the languages in question and about the precise way the theory of grammar should be formulated. (Kiparsky 1982a: vii)

As he has shown repeatedly, rule addition, although possibly the most basic, is not the only mechanism that generative phonology can avail itself of in order to account for phonological change. Kiparsky (1965) notes rule-loss and rule inversion as other possible changes in competence which result in change. Those three mechanisms, together with the assumed changes in representations themselves, can provide an account of numerous phonological changes. The observation that representations also change, and not only rules, is traced by Kiparsky to Hermann Paul, so it was already assumed by the Neogrammarians. It is also recognized by OT, as the already mentioned Lexicon Optimization theorem. In that sense, *SPE* seems like an outlier in wanting to keep representations invariant for long stretches of historical time, and placing the burden of accounting for a change in surface forms on the system of rules only.

Summing up, the idea that lexical representations remain unchanged for long stretches of historical time, which was suggested by early generative grammar is not taken up here. The notion of the coherence of the GVS due to the structural relatedness of its individual shifts is accepted, though, as is Kiparsky's notion that change and structure are intimately related.

2.4.5 OT

OT seems to have, generally speaking, something of an uneasy relationship with the issue of sound change. Just like with the previous generative accounts, the question of change does not figure prominently in the foundational contributions, which concern themselves exclusively with modeling particular synchronic stages of languages. The analogy goes further than that, however, in that just as immediately around the time when phonology was modeled in classic generative grammar, attempts to extend its insights to questions of change began sprouting up (Kiparsky 1965), so with the publication of Prince and Smolensky (1993 [2002]), applications of OT to questions of language change were attempted (e.g. (Jacobs 1995), and (Zubritskaya 1997; Green 1998, both cited in McMahon 2000)).

2.4.5.1 Classical OT

One of the most obvious ways of modeling sound change in OT seems to be constraint reranking. Differences between grammars, be it between grammars of various languages or grammars of various historical stages, can be accounted for in OT by differences in terms of how universal constraints are ranked in a particular language or at a particular language stage. As Prince and Smolensky (1993 [2002]: 6) put it, “interlinguistic differences arise from the permutations of constraint-ranking; typology is the study of the range of systems that reranking permits”. Consequently, it would seem that change can be straightforwardly modeled by reranking constraints. This has indeed been standard practice in accounts of sound change in OT, starting with the very first attempts to do so. And so Jacobs (1995) models a change in syllable structure between Gallo-Romance and Old French (which was linked to the change in stress placement) as a reranking of Alignment constraints. The main advantage she claims an OT account has over previous, rule-based accounts is that a change in syllable structure does not have to be an all-or-nothing affair. If the change consists in parameter resetting, then two polar opposites are allowed in terms of the relation between stress placement (initial or final) and the parameter setting for the mapping of phonological and syntactic boundaries (left or right), with the parameter setting as ‘right’ co-occurring with final stress and the setting as ‘left’ co-occurring with initial stress. However, such an account discounts the possibility of representing intermediate stages of development. According to Jacobs, these can easily be represented in OT, by making recourse to the very core of OT that is to the violability of constraints (1995). Presumably, a constraint can start to make its presence felt by rising in hierarchy, but allow for the existence of forms violating it as long as it is still dominating by some conflicting constraints. The promotion of a constraint in question only results in an overarching change once it sits at a very high position in the hierarchy indeed.

Despite this and other applications of OT to historical developments, however, Kager’s (1999) in many respects comprehensive presentation of the key issues in OT, including problematic aspects and open questions, is completely silent with regard to sound change, or change in general, and these issues are not so much as mentioned. This omission could be seen as telling, as McMahon (2000: 231) notes. It could be telling either of the desire to

forget about diachrony or of the prevalence of more pressing issues. In the hopes that the latter is the case, McMahon (2000: 232) does not fail to point out that

[t]here are good philosophical reasons for considering sound change as necessarily falling within the remit of any decent phonological theory: for instance, the categories of ‘sound change’ and ‘synchronic phonological process’ overlap in membership very considerably, as in the cases of vowel shifts, metatheses, and apparent segmental insertions and deletions; and there is the sheer practical problem of distinguishing synchrony and diachrony, since variation can equally be seen as a consequence of change in progress.

The situation seems to have changed only slightly, as McCarthy’s (2002) *Thematic guide to Optimality Theory* devotes no more than 3 pages to the question of variation and change, and his handbook *Doing optimality theory* (2008), even though it includes the discussion of such a, traditionally, extragrammatical issue as language variation in a section titled “Some current research questions”, goes back to glossing over change completely. It would seem, then, that even though diachronic research in OT has been taking place, it has yet to secure a place in mainstream treatments of OT.

Still, diachronic considerations have important ramifications for the conception of the overall architecture of OT. For one, the way in which language change is accounted for in OT has direct relevance for the issue of the innateness of constraints. This is so, since various historical developments require seemingly language specific constraints to be modeled adequately. If constraint reranking, rather than reformulation, loss or addition is the only way of modeling change in OT, then an account of the loss of a segment or a segment combination needs to assume the promotion of a relevant markedness constraint against that particular segment or segment combination. Given the universal nature of constraints assumed by classic OT, the universal inventory of constraints must include a markedness constraint against each existing segment as well as against each possible segment sequence. This expands the inventory of constraints considerably. One way of avoiding this conundrum is to allow for language specific constraints, that is constraints constructed by children during language acquisition. Then, the number of innate constraints can be kept at a reasonable minimum, but new issues arise, such as the issue of the interaction between innate and learned constraints, as McMahon notes (2000: 233). At any rate, diachronic developments have to be accommodated within any model of phonology, and the way in which such an

accommodation is construed in OT will have far-reaching consequences for such fundamental issues like the relationship between innateness and language-specificity.

The substitution of the rules of generative grammar with output-oriented constraints of OT has thus necessarily led to a change in the mechanisms of change from changes in the rule system of the former to changes in constraint ranking in the latter, with the basic premise remaining the same, that is that change in language equals change in grammars. One point of divergence between OT and SPE, though, is the issue of the historical stability of underlying representations. The introduction of faithfulness constraints in OT, which results in the maximal possible similarity between the input and the output given the markedness requirements, differentiates it from classic generative grammar in that OT does not allow such a great discrepancy between input and output forms as postulated by the SPE. Such discrepancy can arise, but it is not stable, as is the case for example in the account of the GVS proposed by Miglio and Morén (2003), as recounted below.

A model of the GVS in OT which exemplifies the issues presented above was proposed by Miglio and Morén (2003). Theirs is a scenario consisting of three stages. For the first stage, they assume an initial constraint ranking such that vowel length was non-distinctive. All vowels surfaced as long when they were in open monosyllables and monosyllables closed by single consonants, and non-high vowels surfaced as long in open stressed penults. The second stage, the crucial stage of the GVS in their terms, consists in such a reranking of the constraints that the long lax mid vowels became disfavored, and the remaining vowels were raised or diphthongized. These changes pertained to the output only, with no changes in the lexicon yet. The third stage consists in the re-structuring of the lexicon in the next generation of speakers, so that a less opaque correspondence between input and output is achieved.

Miglio and Morén (2003) base their model on the assumption that the pre-GVS quantity adjustments have made vowel length completely predictable. However, the lexical classes distinguished by length before quantity adjustments took place did not collapse altogether. Hence, they would then have to assume that contrasts neutralized in the output were maintained in the input, which is untenable as due to lexical optimization the neutralized contrasts would have disappeared from the input as well. This does not seem to be what they assume. Their position, instead, rather seems to be to assume that all inputs were short. But

this would mean the collapse of the contrasts and is therefore untenable as well. The view subscribed to in the account presented in the present thesis agrees as to the importance of the weakening of length as a feature distinguishing vowel contrasts with Miglio and Morén's (2003) account. However, it does not assume a stage at which the contrasts were abolished altogether. Rather, the weakening of length as a feature linked to the growing entanglement of its phonetic correlate, i.e. duration in the expression of other factors besides length, is assumed to have allowed the rise of primarily quality-based tense/lax contrasts, and the weakened role of length in English phonology is held responsible for the instability of the vocalic system to this day.

Another assumption of Miglio and Morén's (2003) model which can be criticized is their treatment of the GVS as a synchronic shift. To be able to see the GVS as a synchronic shift, they are forced to assume in their three-stage model the existence of Stage 2, a stage at which speakers have the underlying forms identical to those of speakers at Stage 1, and output forms identical to those of speakers of Stage 3. The existence of speakers who posit the same input as speakers at Stage 1, but a completely different constraint ranking (which leads to different surface forms), is, however, highly problematic. If the ambient language causes Stage 2 speakers to have an innovative constraint ranking, then they would not have posited the same input as previous speakers. And if the ambient language prompted them to posit the same input as that of previous speakers, then they would also have posited the same constraint ranking.

Further, in the details of their analysis, where the pre-GVS quantity adjustments crucial to their account, as well as the Stage 2 of the GVS itself, change in surface forms is 'explained' as taking place because of a particular instance of constraint reranking. The reasons for the reranking, however, are often not presented. With regard to one of the quantity adjustments, for instance, they maintain that "[l]oss of distinctive vowel length before non-geminate codas is due to [the following] ranking [...]" (Miglio & Morén 2003: 198). Thus, the reranking is given as a cause of change, while the reranking itself is not given explicit motivation. If the motivation were the suspension of vowel length distinctions, which could be assumed due to Miglio and Morén's (2003: 196) subscription to the 'holistic' view of quantity adjustments, then the reranking is motivated by its goal. Such teleological motivation is, however, clearly not a welcome result, as such teleology would have to assume

speakers actively directing change, with all the problems involved in assuming human agentivity in change. Here, speakers would have to know that a particular constraint ranking is overall better than another, and decide to construct a better ranking, one that does not correspond to the linguistic data they are faced with, but one that they deem better.

Aside from the implicit teleology of change conceived of in terms of constraint reranking, such treatments do not go far beyond merely describing, rather than explaining change, as argued in a critical assessment found in McMahon (2000). The example she refers to is that of the loss of the /kn/ cluster from onsets of English syllables. She refers to an analysis by Green (1998, cited in McMahon 2000) under which this change came about by the markedness constraint against this cluster - *a[kn – being promoted over the faithfulness constraint which militates against deletion – MAXIO. The question of why the constraint was promoted, however, remains unaddressed. If the difference between the constraint ranking assumed for one language stage differs from the constraint ranking of another because of the different grammars constructed by language learners during language acquisition due to lack of robust evidence to the contrary in the environment (Archangeli 1997: 31, referred to in McMahon 2000: 234), then the first element of the cluster must be assumed to have disappeared before the constraint reranking took place. Consequently, reranking is merely a description of what happened, with no claim to causality whatsoever.²⁰

2.4.5.2 The life-cycle of phonological processes and Stratal OT

McMahon's, as well as Kiparsky's (cf. Section 2.4.4 above) views on the importance of diachronic explanations for phonology are shared by Bermúdez-Otero (2013 (forthcoming)), who takes up Kiparsky's (2006) idea of 'amphichronic' phonology, that is of a phonological explanation that draws on both diachrony and synchrony. Due to the 'overlap in membership' between instances of sound change and synchronic phonological processes mentioned above, one of the goals of phonological research is to identify those properties of sound systems which exist only because of the historical trajectories taken by given languages, and not because of constraints on the architecture of synchronic grammars. As a step towards a demarcation between those characteristics of a phonology which are due to the way

²⁰ For other criticisms of the adoption of constraint reranking as an explanatory mechanism of language change in OT, see also: Gess (2003), McMahon (2003) and Reiss (2003).

phonology operates and those which are only remnants of historical developments, Bermúdez-Otero and Trousdale (2012) suggest the idea of a ‘life-cycle’ of phonological processes (reproduced below as Figure 3).

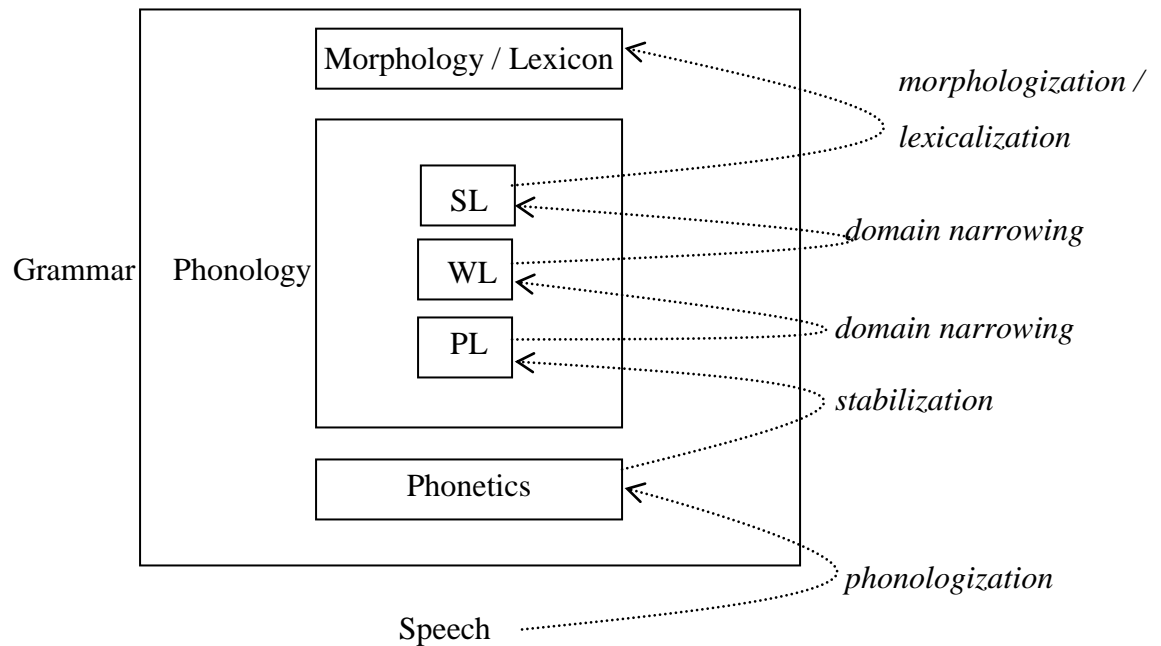


Figure 3: The life-cycle of phonological processes, reproduced from Bermúdez-Otero and Trousdale (2012: 700).

SL = stem level, WL = word level, PL = phrase level

The general idea of a life-cycle of a phonological process is that patterns present in speech, through phonologization, become part of (language-specific) phonetics as gradient processes. This happens when speakers misinterpret unintended characteristics of the speech stream as intended, that is as grammatical. Through stabilization, a process enters the phonological component. At that point, a gradient phonetic process turns into categorical phonological processes, and it becomes active on the post-lexical level. Through successive cases of domain narrowing, it succeeds to act first on the word level, and then on the stem level. The last stage of the life-cycle is for a process to undergo morphologization or lexicalization, when it becomes part of the morphology or the lexicon (2012).

A promotion of a process one level up in this hierarchy does not have to mean that it ceases to operate at the lower level; for example, in English, stem-level categorical

palatalization as in *confess* /kən'fes/ ~ *confession* /kən'feʃn/ coexists with gradient, phonetic palatalization as in *press you* /'presʃu:/ (Bermúdez-Otero 2007: 506). Such a co-existence of similar, or rather, diachronically related, processes on different levels was dubbed 'rule scattering' by Bermúdez-Otero (2013 (forthcoming)). Another example from PDE that Bermúdez-Otero (2013 (forthcoming))§§ 2.3) gives is that of short *a*-tensing in Philadelphia. Short *a*-tensing is a gradient phonetic rule in Philadelphia, subject to numerous aspects of the consonantal environment, such as, notably, presence of nasals in the coda. However, at the same time, short *a*-tensing is operative also at the stem-level, where the default distribution of tense and lax /æ/s is governed by a categorical stem-level rule, expressed as a stem-level constraint ranking in Bermúdez-Otero (2007§§ 21.3.3).

Naturally, the assumption behind the life-cycle model is that, in synchronic phonologies, phonological processes apply at various derivational levels. This was the insight of lexical phonology (Kiparsky 1982b), which constrained the unlimited nature of rule ordering of classic generative phonology to three levels, namely the stem level, the word level, and the post-lexical level. Though there was no place for such a division in the early installments of OT, due its strict parallelism, it has been married with OT as 'Stratal OT' (Bermúdez-Otero 2007) or 'LPM-OT' (Lexical phonology and morphology OT) (Kiparsky 2000). In these incarnations, the different application of phonological processes at the three levels translates into differing constraint rankings applying at the three levels, with the output of the stem level being the input to the word level, and with the output of the word level being the input to the phrase, or post-lexical level. In keeping with Kiparsky's terminology, phonology beyond the word level is referred to as 'post-lexical', rather than phrase-level phonology.

For reasons given elsewhere, in the present thesis, the formalism of OT is adopted. OT tableaux are employed to represent the grammars of the successive stages in the development of the English vocalic system. Following the objections to constraint reranking as a mechanism of change, sound change is not conceived of here as taking place solely as reranking, but also as reformulation of constraints. This rests on the constructionist, rather than innatist view of constraints. The concept of the life-cycle of phonological processes is adopted, which implicates a stratal, rather than strictly parallelist version of OT. To overcome the weakness of the lack of explanatory power of assuming that changes in

constraint rankings represent all that is interesting, the influence of language use on language change is admitted through the incorporation of the insights of the exemplar-based phonology, to be described below.

2.4.6 Exemplar-based phonology

The last two decades have seen the rise of exemplar-based approaches to phonology (e.g. Miller 1994, Johnson 1997, Bybee 2001, Pierrehumbert 2002b, Phillips 2006). Though relatively recent as a serious contender in modern phonological theorizing, a frequency-based view of phonology can trace its origins already in the writings of nineteenth century linguists. As already noted in 2.4.2, Hermann Paul's views on the representation of sounds assumes something like multiple memory traces of the same sound. The second important precursor of exemplar theories is the already mentioned Hugo Schuchardt (1885), who, criticizing the Neogrammarian view of sound change as exceptionless points out that individual lexical items undergo sound change at different times, and that, far from being random, their likelihood to participate in a sound change depends on their frequency of use, in that "[s]ehr selten gebrauchte Wörter bleiben zurück, sehr häufig gebrauchte eilen voran" (Schuchardt 1885: 25). According to him, sound change starts with the most frequent words because, on the one hand, speakers have the greatest need to make pronunciation easier in the items that are used most often, and, on the other hand, because the most frequent words carry the least risk of being misunderstood (1885: 27). His own proposal of studying sound change does not go far beyond stating that frequency of words can lead to sporadic sound changes and that the intra- and interspeaker variation cannot be left out of linguistic accounts.²¹ Still, by demanding that the implications of this observation be accounted for in an adequate theory of change, he started a mode of thinking that ran parallel to the Neogrammarian, and later structuralist and generativist view of sound change (Murray 2009).

The most comprehensive presentation of a model of phonology nested in this tradition is Bybee's (2001) *Phonology and language use*. Therein, she puts forth a theory unifying language substance (phonetics) and language use as dominant in shaping language structure. In a reversal of priorities, where following the separation of 'langue' and 'parole' or of

²¹ Recently, Phillips (2006: Chapter 3) takes Schuchardt's idea as a starting point of a more elaborate theory of change.

competence and performance it was the first member of each pair that deserved linguists' attention, Bybee focuses on language use, and sees structure as 'emergent'. One of the main links between use and structure is the many manifestations of frequency effects on synchronic grammars, which have direct implications for diachrony. Highly frequent forms are more likely to undergo phonetic change, but are at the same time "more conservative in the face of grammatical change or analogical change based on the analysis of other forms" (Bybee 2001: 12). The important level of knowledge is the lexicon, with individual entries containing fine grained phonetic information fully-specifying the items in question. As such, phonological segments are posited to emerge as generalizations over a number of tokens which contain them.

Despite its intuitive appeal to and possible precursors in the field of historical linguistics, exemplar-based approaches to phonology have some serious shortcomings with regard to accounting for language change. Many of them stem from the rejection of segments as phonological primitives and from encoding phonetic detail directly in lexical representations.

While it might be the strength of Bybee's approach that it is able to deal with isolated, exceptional developments which are consistent with the basic assumptions and mechanisms of frequency-based models, it is at the same the weakness of this model, as pointed out by, for instance, by Bermúdez-Otero (2007), that it in fact predicts that most of sound change takes such a course. If there is no unity among the various instances of phonemes, than it is the long-term stability of many such entities that poses a problem. The assumption that each lexical entry is connected to its own exemplar cloud, in view of the fact that most diffusing changes stop before they affect the entire lexicon (Wang 1969), leads to some unlikely implications for sound change (Bermúdez-Otero 2007: 513). First, 'endogenous', that is variety-internal, lexical splits should be commonplace. Lexical splits are, however, very rare as such, and the instances that have been observed seem to have been caused by language contact rather than arisen endogenously. Second, the lexicon should be full of remains of unfinished diffusing changes. Even though the remains of some arrested changes can be found²², it is rather clear that such cases are not widespread, that is that the phonologies of

²² Such as English *ass* and *bass*, which took part in early, sporadic r-deletion.

known languages are not littered with holdovers from the past. Hence, the issue of stability of phonemes over time is what poses difficulty for exemplar-based accounts.

Realizing the need for phonemes, or “units of sound structure”, whose existence is clear not only from historical data but also evident in speech processing in the pronunciation of neologisms and loan word assimilations, Pierrehumbert (2001: 139) stresses that “the correct model [of phonological competence KK] must describe the interaction of word-specific phonetic detail with more general principles of phonological structure”. The problems raised for exemplar-based theory by the issue of phonetic residue can be dealt with by unifying exemplar based approaches with formal models. Such approaches have come to be known as ‘hybrid models’, and a number of them have been proposed (e.g. Hawkins 2003, 2010; McLennan & Luce 2005; Luce & McLennan 2005.; Pierrehumbert 2006, for an overview of experimental evidence for those cf. Nguyen 2012). Pierrehumbert (2002b: 107), for instance, suggests that, instead of being fully phonetically specified, lexical entries could be enriched with a numerical index of lexical accessibility, and that the phonetic effects of a phonetically-gradient pronunciation could be derived from that. This approach is taken up in Chapter 4 below.

Another problem with the exemplar-based account of Bybee (Bybee 2001) is the issue of the role of ‘similarity’ in the model of phonology she puts forward. Silverman (2001) raises this issue as a case of lack of terminological clarity in Bybee’s book, but in fact this issue goes beyond terminological imprecision. It is key for Bybee’s model to work that ‘similar’ portions of the speech signal be considered as instantiations of the same category, once the category has emerged. Now, it is crucial for the discussion of sound change to define clearly how this similarity is measured and how different do sounds have to be for two separate categories to emerge. In fact, since any two productions of the same sound are different, and since the productions of a ‘similar’ sound in different phonetic environments can be very disparate, splits should be expected to happen constantly.

The view of the language structure as emergent from language use is reminiscent of Skinner’s behaviorism, and shares with it an important weakness. Though there are important differences between the two approaches, such as that it is reinforcement that is the main force behind linguistic structure in behaviorism and it is various aspects of language use, such as

frequency of use, ease of processing and so on in usage-based phonology, both are similarly evasive when it comes to dealing with the issue of creativity or novelty in language. Skinner (1957) ascribes these abilities to ‘generalization’, and Bybee (2001: 15) states that “the repetition of gestures and sequences across words allows relations of identity and similarity to develop in stretches of speech, giving rise to segment, syllable and foot-sized units”. The use of generalization as a non-telling cover term for the essential quality of human language is, of course, what Chomsky (1959) criticized vehemently as a major flaw of behaviorism, and it is the focus on how this ‘generalization’ takes place that launched the generative enterprise. Bybee’s proposal is very different from Skinner’s, but it could be argued that the role of ‘schemas’ is reminiscent to Skinner’s ‘generalization’. Among the differences is the fact that Bybee (2001) does not share the view of behaviorists that what goes on in the minds of speakers is beyond scientific investigation. Further, she recognizes that even though much of what people say may consist largely of pre-fabricated chunks of what they have repeatedly heard and said before (Erman & Warren 1999, cited in Bybee 2001: 15), novelty is a fact of language and cannot be swept under the carpet. She bases her account on recent findings regarding the storage of lexical units, and proposes that lexicon is characterized by organized storage, where each word is tightly embedded in a network of connections, where similar words are linked to each other, and where similar parts of words are linked to each other. The connections between the similar components act as ‘schemas’, or generalizations, which are argued to better capture the cognitive reality of linguistic knowledge of speakers than symbolic rules do. One crucial difference is supposed to be that schemas, as opposed to rules, do not have an independent existence separate from the individual items they link; they are “non-process statements about stored items” (Bybee 2001: 22). This intimate linking of schemas to individuals entities, however, does not logically exclude the possibility that they (categories and rules) exist, to a certain extent, ‘independently’, and that no reference to the generalization itself can be made as opposed to all of its individual instantiations. For all those differences, then, the similarity between the approaches of Skinner and Bybee lies in the fact that by focusing on what is, according to them, worthy of study (acquisition through reinforcement, for Skinner, and the manifestation of frequency effects on language structure for Bybee), they downplay the role of grammar. It seems that it is the need to stress the hitherto often ignored aspects of linguistic knowledge, rather than these findings themselves

that lead to the radical rejection of the independence of categories and rules. In a way, just as behaviorists pointed out important aspects of linguistic behavior, and as imitation and (broadly construed) reinforcement do play a role in language acquisition, as explicitly advocated by the evolutionary approach, similarly the role of frequency effects cannot be denied; still, Bybee (2001: 13) warns about the limits in the degree of similarity between the generalizations devised by speakers and those “devised by linguists on the basis of distributional evidence”. Since it is distribution alone that speakers go by, though, it is unclear why linguists should fail at arriving at just the right generalizations.

Overall, the main claim of the exemplar theory, namely that language use has influence on language structure, is valid and needs to be incorporated within a linguistic framework which strives to take into account the cognitive reality of speech processing. Even if phonetically rich tokens are stored in memory, however, the reference to exemplars does not obviate abstract units such as phonemes. Therefore, a phonological model incorporating both usage-based frequency information and the existence of formal categories, along the lines of the hybrid models mentioned above, is advocated.

2.4.7 Summary

To summarize the discussion in Chapter **Error! Reference source not found.** so far, which consisted in the evaluation of the contributions of the Neogrammarians, structuralists/functionalists, generative grammar, OT, and exemplar theory to the study of language change, the following view of language and of language change emerges.

The entities undergoing sound change are lineages of successive mental representations of sounds. These are implemented as neural patterns in the brains of speakers. For example, one such lineage is that of the mental instructions which resulted in the successive pronunciations of the stressed vowel of PDE *make*, namely: /ɑ > a > a: > æ: > ɛ: > e: > eɪ/. Such lineages lie at the center of this perspective, since it is believed that the way in which they are transmitted, i.e. diachrony, is important for the understanding of the systems which they form, i.e. for synchrony.

The most relevant kind of change is change happening below the level of consciousness of speakers, and their role is, for the most part, left out of the discussion. It is acknowledged that in certain individual instances the rate of change can be halted or, possibly a certain change can be initiated as a result of language planning, but it is unlikely to be the case with sound change. Even when language planning were brought into the picture, the ontology of language advocated here would not have to be abandoned. Just like pigeon breeders purposefully selecting pigeons with desired traits and eliminating those with undesired traits from the population can be thought of as an environment, albeit a very special one, in which pigeons evolve, so can be willful efforts of language planners be thought of as environment in which language evolution unfolds.

As for the role of children versus adults in language change, in line with the Neogrammarian position, both possibilities are admitted, with the entrenchment of frequently recurring speech patterns exerting influence on the kind of input that forms the basis for the language acquisition of future generations. Language change is instantiated both through first language acquisition and throughout life. As competence constituents are transmitted from speakers to children acquiring language, there is a great chance that the way in which individual competence constituents organize to form a linguistic system will show considerable differences to the original ones. At the same time, it is possible for change to continue throughout life, but these kinds of changes are expected to be less fundamental.

Preservation of contrasts is an active force in language change, and it forms part of synchronic grammars in the form of PC constraints. Minimization of effort might also play a role in sound change, e.g. in the development of historical vowel reduction and loss. Crucially, the effects of these two kinds of factors manifest themselves because they constitute pressures influencing the process of the replication of linguistic constituents, and not because of the will of the speakers.

Language change, as an instance of evolutionary change, is adaptive. Hence, an innovative structure must have certain advantages increasing the rate of its replication over the current structure if it is supposed to replace it. However, the fact that a certain new feature has a replicative advantages over another feature does not mean that the new system is overall better than the old system.

Lexical representations are close to surface forms, and so a case of language change in which surface forms are observed to have changed involves changes in the representations. The structural relatedness of its individual shifts involved in the GVS makes it a coherent event.

Synchronic stages of languages represent temporary systems consisting of linguistic replicators (these include lexical representations, phonemes that these are composed of, as well as competence constituents for units of rhythmic organization, i.e. feet) as well as of the grammar, that is of a conventional resolution of the conflicting pressures on the expression of the replicators. In the present thesis, the formalism of OT is adopted to present grammars as constraint rankings. Following the objections to constraint reranking as a mechanism of change, sound change is not conceived of here as taking place solely as reranking, but also as reformulation of constraints. This rests on the constructionist, rather than innatist view of constraints (discussed in greater detail in the next section). If, at least some, constraints are thought of as changing, then they would also be replicators. The concept of the life-cycle of phonological processes is adopted, which implicates a stratal, rather than strictly parallelist version of OT.

To overcome the weakness of the lack of explanatory power of assuming that changes in constraint rankings represent all that is interesting, the influence of language use on language change is admitted through the incorporation of the insights of the exemplar-based phonology. However, the reference to exemplars does not obviate abstract units such as phonemes. Therefore, a phonological model incorporating both usage-based frequency information and the existence of formal categories, along the lines of the hybrid models is advocated.

2.5 FORMALIZING EVOLUTIONARY LINGUISTICS IN OT

OT is a good choice for presenting the proposed mechanism of vowel shifts, since it comes with a range of analytical tools suited for the representation of the various synchronic stages of the language. Research in OT, which has been going on for over two decades now, has focused mostly on phonology and it has resulted in the discovery of a number of constraints and their interactions as far as the phenomena currently under investigation are concerned.

With reference to the issue investigated in the present thesis, the OT approach of presenting how conflict between various rhythmic constraints is resolved is particularly useful. In particular, the constraints whose reformulation is thought to have played a role in [\pm tense] taking over [\pm long] as the primary feature distinguishing between two classes of vowels in English, a development crucial in the account to be proposed, are the Preserve Contrast constraints.

With regard to phonological change, as sketched in Section 2.4.5, researchers have applied OT in diachronic investigations. Previous treatments of language change, chain shifts, and the Great Vowel Shift specifically have provided a number of analytical tools to model these phenomena. With all the benefits of the OT approach, though, in view of its weaknesses when dealing with change, placing it within evolutionary linguistics and thus adding the population perspective is thought to be a welcome extension.

The insights gained from the overview of findings of the various research traditions, an overview performed from the perspective of evolutionary linguistics, form the conceptual toolkit applied in the present thesis. The account proposed here is then formalized in OT. A presentation of the basic architecture of OT has already been provided, as well as the treatment of language change in classical OT and in Stratal OT. The choice of Stratal OT over strictly parallel OT was dictated in part by the evolutionary agenda of the present thesis, in that the recognition of the importance of the historical paths of replication for synchronic sound patterns of languages, a hallmark of the evolutionary perspective, finds a natural ally in the idea of a life-cycle of phonological processes, a concept much more easily operationalized in Stratal, than in parallel OT. The fact that OT is used here to formalize claims made from the evolutionary perspective results in a stance on two further issues, namely the issue of the innateness of constraints and the issue of language acquisition, which are different to views of classical OT. As noted in Section 2.4.5, OT attempts at modeling phonological change have been shown to have some inherent problems. Thus, it should be welcome, speaking from an OT perspective, if these could be overcome when placed within the framework of evolutionary linguistics.

2.5.1 Some modifications to mainstream OT stemming from its combination with evolutionary linguistics

First of all, accounting for the learnability of OT grammars which correspond to the cognitive reality of language acquisition requires that phonemes be separate from constraint rankings which ‘produce’ them. Under classic OT conception, the set of phonemes manifest in a given language or at a particular language stage follows from the ranking of the relevant constraints. If a given language is characterized by a particularly highly ranked markedness constraint against, say, voiced consonants, with no concomitant context-specific faithfulness constraint requiring the preservation of voicing in consonants in some specifiable environment, then voiced consonants will not surface at all. Should they have voiced consonants underlyingly, there will be no evidence of that in the output. Hence, under Lexical Optimization, speakers constructing their grammars when exposed to the outputs of the grammars with this ranking, will not have voiced consonant phonemes. When this logic is followed strictly, then phonemes are not granted existence separate from the constraint rankings that they follow from.

This conception, however, is inconsistent with a cognitively plausible view of language acquisition. In first language acquisition, children have the double task of (a) acquiring the lexicon and (b) acquiring the constraint ranking (for an overview of the topic see: Kager [1999: 296]; McCarthy [2002: 202], [2008: 264]). The mechanism proposed for the formulation of the constraint ranking is that of Constraint Demotion (Tesar & Smolensky 1993). The basic idea is that children start by ranking markedness constraints above faithfulness constraints, and arrive at an adult-like ranking by demoting constraints for which they encounter evidence pointing to their lower position in the hierarchy.²³ The mechanism of Recursive Demotion, together with the assumption about the initial state, has been shown

²³ One of the crucial assumptions for this mechanism to yield the right ranking is to posit that children start off by ranking markedness constraints over faithfulness constraints. They then, through exposure to ambient language which has the marked structures, that is to speech produced by speakers in whose grammars some faithfulness constraints outrank the relevant markedness constraints, come to demote them. The assumption that in the initial state of the grammar markedness constraints outrank faithfulness constraints is crucial. Without it, learners would never arrive at a ranking which bans the sorts of structures forbidden in the ambient language, as positive evidence alone does not suffice to do so. This crucial assumption has its predecessor in the hypotheses of Natural Phonology (Stampe 1979), where children are posited to start off with a number of rules that make articulation easier (called ‘natural processes’), which have to be suppressed for the child to arrive at adult-like phonology.

to be able to make the right predictions about the constraint ranking. However, one of the assumptions it incorporates is that the learners possess an adult-like lexicon from the start, it is the sort of idealization that has to be done away with for the modeling to resemble the acquisition process, since it is an essential component of the task that children do. As Kager (1999: 322) points out, the living child does not have adult input forms, but only ‘hypothetical inputs’, which have to be monitored as the acquisition process proceeds. Essentially, three variables must be manipulated by children at the same time, i.e. (a) the constraint hierarchy (b) output representations and (c) the underlying representations. Kager (1999: 323) suggests that Tesar’s (1996) extended model is capable of incorporating all three of them. It seems to be computationally viable to assume a cyclical process of acquisition, where acquiring a portion of lexicon, followed by a partial ranking of constraints in turn leads by approximation to the mature state of the grammar. When this is related to the cognitive reality of language acquisition (as shown by the research in exemplar theories, cf. Section 2.4.6), then these hypothetical representations are drawn from a pool of multiple tokens stored in the memory. If that is the case, then it must be admitted that structural elements, i.e. lexicon entries but also phonemes, can be perceived and stored which would not be admitted by the particular constraint ranking at the moment of their acquisition, but which later become admissible when the ranking is restructured in the next step in the cycle. If it is admitted that constraints are reranked to better fit the phonemes a speaker has, then the phonemes must be granted existence separate from the constraint ranking itself. As such, the existence of phonemes is not only an empirical finding, but also a necessary assumption for the formal model to work. If that were not the case, then the whole acquisition process would be thwarted.

Even though the plasticity of all components of the linguistic competence diminishes with age, as does plasticity in other cognitive and bodily domains, it does not disappear completely. The possibility of phonetic change throughout a person’s lifetime has, for example, been demonstrated by Harrington (2007). Under OT such change must arguably mean a restructuring of the grammar. The hypothetical case of an individual who leads a large part of their life with lax /ɪ/s word finally, with pronunciations such as *happy* /^hæpɪ/ or *very* /^lveri/, and who comes to pronounce these words consistently as /^hæpi/ and /^lveri/ later

on, can illustrate of the point. A change of this sort, interpreted as a change in grammar, amounts to banning this lax /ɪ/ from final position. Now, the introduction, or promotion of the constraint on /ɪ/ in final position must, or so theory has it, be deduced from memorized

representations of the relevant words, with the new value of the phoneme occurring in this position. Since adults normally do not acquire vast amounts of new lexical items on the basis of which an innovative grammar would be structured, it must be assumed that they memorize the new phonemic representations in addition to the ones they have already stored. Hence, the phonemes of the new /'hæpi/ and /'veri/ representations must come to exist independently of, and unaffected by the originally acquired constraint ranking, since they would be ruled out by it if it did not change, but can only cause it to change after having been internalized. This means that also for adults the possibility that phonemes can exist independently of constraint rankings must be posited.

The discussion above has established the argument that phonemes exist independently of constraint rankings. As such, it stakes out a claim about the nature of phonemes which can be situated within the broader discussion of phonemic representations. As far as the representation of phonemes is concerned, two questions have to be answered, which are the following: first, what kind of an entity phonemes are, and second, what their content is. With regard to the issue of the ontology of phonemes, Twaddell's (1935) monograph outlines the possibilities available to phonological analysis. One position which he severely criticizes is one assuming that phonemes are a physical reality, present directly in the speech stream. This was the position of structuralists, e.g. Bloomfield (1933), who claimed that phonemes could be recoverable from the acoustic speech stream. The failure (Dresher 2011: 244) of the attempts at identifying invariant acoustic correlates of phonemes or of features in the speech stream, however, shows that Twaddell's (1935) criticism of this conception was well-founded. Another position on the issue of the mode of existence of phonemes is that they have psychological reality (Sapir 1925; Trubetzkoy 1939). Twaddell rejects it on the same grounds Bloomfield did, namely because of the impossibility of a scientifically founded study of the postulated mental objects. Twaddell's (1935) conclusion from the flaws of these two, untenable in his view, positions is that phonemes are creations of analysts, with no reality for language users themselves. This is, however, not the only possible conclusion. As

pointed out by Dresher (2011: 245), the belief that mental constructs are not explanations but rather the things to be explained has resulted in a good deal of fruitful research following the cognitive turn. Additionally, with the rise of neuroscience, methods are becoming available which will make possible the investigation of these, hitherto largely unknown, mental entities. What follows from this is that they can be conceived of as having physical representation in the brain. This is what lies behind the conception of phonemes in Ritt (2004). Phonemes, as mentioned in the outline of the evolutionary approach in 2.2 above, are assumed there to be physically manifested as neurally implemented associative patterns. They are defined as very much real entities, as patterns unifying auditory impressions with articulatory instructions.

To turn to the question of the content of phonemes, three different possible positions can be identified in the literature, as suggested by Anderson (1985) and taken over by Dresher (2011). All three of them are, according to Anderson (1985) different possible interpretations of de Saussure's conception of the linguistic sign. First, there is the notion, associated with the conceptions of Trubetzkoy and Jakobson, that phonemes contain only the distinctive information needed to distinguish between various phonemes of the same language, with all predictable information being relegated to a separate component of grammar, dealing with phonetic implementation. This means that the basic variant of a phoneme is "incompletely specified", and that its allophones come into being by filling in all the predictable information. In opposition to that view is the "fully specified basic variant" theory, which according to Dresher (2011: 250) is similar to that adopted by some proponents of Cognitive Grammar, and according to which there is one basic variant, with all its details stored in the representation, and all its allophones are the result of some kind of transformations of it or other. The last possibility identified by Anderson (1985) is that of a "fully specified surface variant", where each allophone has its own representation, and where the grouping of allophones into a phoneme arises through the rules or constraints of the grammar, and not through any intrinsic feature of the representations²⁴. As Dresher argues, this view essentially

²⁴ This would be compatible with the way in which aspirated and unaspirated voiceless stops come to be complementarily distributed in English under the analysis provided in Section 2.3.2 above to illustrate Lexicon Optimization. It is simply through the ranking in which the context-sensitive markedness constraint (in case of aspiration: *^l[-voice, +stop]V) dominates the context-free markedness constraint (*[+aspirated]), which in turn dominates the faithfulness constraint (IDENTIO(aspiration)) that these stops are complementarily distributed,

obviates the need for phonemes. Although he equates this last position with “some versions of Optimality Theory” (Dresher 2011: 250), I argue here that this is not a necessary pairing, and that in a more cognitively realistic take on OT, one that acknowledges that languages are acquired, and that takes into account the findings of laboratory phonology which posit the need for phonemes (Pierrehumbert 2001: 148), the phonemic level does have its place. Indeed, under Ritt’s (2004) account, where a phoneme is a pattern in an associative network, the boundaries between the first two accounts, the incompletely specified theory and the fully specified basic variant, are rather blurred, since a phoneme, as a neural network, is linked to all the relevant nodes for articulatory gestures and auditory impressions, so there is clearly only one representation, though, since in combination with different positional nodes, such as ‘onset’ or ‘coda’, for example, different components of the network are activated. Thus, based on the arguments provided above in this section (that constraint rankings are arrived at in first language acquisition on the basis on phonemic representations, and that, similarly, grammars can change within the lifetime of an individual on the basis of new phonemic representations) I assume the reality of phonemes conceptualized in the way described above, whose existence is separate from the relevant constraint rankings.

Furthermore, it can be speculated that, contrary to OT’s assumptions of the richness of the base, not all imaginable inputs are submitted to EVAL, but only those consisting of phonemes arrived at by exposure to and experience with the ambient language.

Another consequence of assuming an evolutionary, that is cognitively realistic, perspective for one’s views on OT is that a plausible account of why constraints are functional must assume that constraints are not innate. As McCarthy (2008: 220) notes, OT is a framework which overcomes the weaknesses of purely formal grammars by integrating functional tendencies, and at the same time overcomes the weaknesses of functional approaches in which the adequacy of competing explanations cannot be easily assessed. The functional grounding of violable constraints leads to the integration of functional tendencies into a formal model of grammar. Stressed already in Prince and Smolensky (1993 [2002]: 198), the profitability of incorporating phonetic functionalism into formal phonology is seen in the ‘phonetically driven phonology’ line of research of Hayes and Steriade (2004). With

and can therefore be considered allophones of the same phoneme. Other possible rankings of these three constraints could yield lack of variation, positional neutralization or full contrast (Kager 1999).

regard to taking phonetic functionalism as a criterion in establishing the legitimacy of constraints, Hayes (1996: 291) writes:

In the case of phonetic functionalism, a well-motivated phonological constraint would be one that either renders speech easier to articulate or renders contrasting forms easier to distinguish perceptually. From the functionalist point of view, such constraints are *a priori* plausible, under the reasonable hypothesis that language is a biological system that is designed to perform its job well and efficiently.

This functional justification or grounding of constraints is what provides a direct link to evolutionary linguistics. Under the evolutionary perspective, competence constituents replicate under selective pressures of the environment, part of which is the speaker construed as a biological organism, part of which is its (the organism's) mental life, including the socially embedded needs to belong on the one hand, to assert individuality on the other. Although the latter, which arise from the fact that speakers are parts of social groups are, arguably, not easily handled within OT, the 'needs'²⁵ of the speakers which arise from their physiological make-up can be thought to be reflected in the constraints of OT. Assuming that OT constraints are functional, then, in the sense that they reflect the limitations of articulation and perception as well as the need to maintain contrasts, a question that needs answering is: how have constraints come to reflect functional tendencies? This question is far from trivial, as on its answer hinge fundamental issues relating to innateness and language universals. McCarthy (McCarthy 2008: 222-223/ yearonly) categorizes the kinds of answers to this question which are in circulation into three groups.

First, in line with the formulation of Prince and Smolensky (1993 [2002]), constraints in mainstream OT (including the classic work on learnability in OT, i.e. Tesar & Smolensky 1993) are argued to be innate. If they are posited to be both innate and functional, then to account for their functionality one would have to stipulate that they co-evolved in humans together with the development of the cognitive system, the voice tract and the auditory apparatus, and that this is why they reflect the limitations of all those components. Faced with the multitude of rather parochial-looking constraints, however, it is difficult to see just how these would evolve as a result of biological natural selection. Effectively, one would have to argue for a selective advantage for organisms possessing each and every constraint

²⁵ The use of scare quotes around *needs* is dictated by the fact that these amount to constraints on the replication process, a usage rather different than that employed in speaker-centered accounts.

(McCarthy 2008: 222), including, for instance, all context-free featural markedness constraints and context-specific featural markedness constraints, with both groups of constraints placing partly conflicting requirements on output forms. Every single one of them, with a very large group of (low-ranked) constraints not showing any effects in output forms of many languages, would have to be shown to contribute to their speakers successfully reproducing.

Two further possibilities share the assumption that constraints are not innate but that they are constructed by learners on the basis of ambient language. They differ, however, as to how functionality enters into the picture; the first view being that it reflects patterns of historical change, the second that it reflects the nature of the language acquisition process itself.

The first of those, advocated for instance by Haspelmath (1999), is that learners deduce constraints from ambient language, and the fact that these constraints reflect functional tendencies is ascribed to issues of historical change. The reasoning is that certain types of historical change are more likely to happen than others for functional reasons, which leads to certain structures re-occurring in various languages more often than others, and this, in turn, leads to the construction of constraints which reflect the functional tendencies. The problem with this view is the same as with other ‘diachronic reductionist’ (Bermúdez-Otero 2006) approaches to markedness, namely that the combined result of common functional historical developments could in theory result in unattested synchronic phonological processes, such as final voicing; the nonexistence of such processes is taken as proof of “some constraint on language design (whether structural or functional)” by Kiparsky (2006: 224).

The second non-innatist approach to explaining the functional character of constraints is that learners discover constraints sometime in the babbling stage while experimenting with all possible sounds, and the sounds that cause difficulty lead to the formation of corresponding markedness constraints ((Hayes 1996; Boersma 1998; Becker & Tessier 2011). Hayes (1996), for instance, puts forward a formal account of a mechanism of ‘inductive grounding’, which could account for the way in which learners construct phonetically grounded constraints. In short, it is assumed that learners, based on their experience with their own articulation and perception, set up ‘phonetic maps’ reflecting the difficulty of particular sounds, and that they, on the basis of these maps, prolifically construct

constraints²⁶. These constraints are then fed into the Constraint Demotion algorithm of Tesar and Smolensky (1993) referred to above, with a large number of constraints ranked so low and therefore not having empirical results.

McCarthy (2008: 223) raises the objection to Hayes' (1996) account by noting that speakers possess constraints for which there is no evidence in the ambient language, such as the constraint against voiced geminates possessed by speakers of a language in which these do not surface at all. The construction of constraints for which there is no evidence in ambient language, however, follows from the very basic assumption that children set up constraints in response to the difficulty they encounter when experimenting during the babbling stage.²⁷ The view that constraints are constructed by learners thus is the most plausible of the possible consequences of the assumption that constraints are functional.

Supportive of the constructionist view of constraints are also the results of research into transient phonology. McAllister Byun et al. (2012) persuasively show that to account for the uniqueness of processes present in child language which are absent from any known languages, one needs to assume the existence of child-language specific constraints, which are thrown out of speakers' constraint sets as maturation proceeds. Such child-phonology specific constraints are set up in response to phonetic difficulty which corresponds to the developmental stage the children are at, and so they clearly are discovered, and not part of a universal inventory. Thus, once it is acknowledged that constraints can enter the constraint set by discovery when it comes to child phonology, it is not far-fetched to assume that other constraints, those that do form part of adult-state grammars, are also constructed, rather than innate.

Some of the already mentioned selective forces that provide direction to language change are exactly the sorts of constraints that form part of the OT framework. Markedness constraints, which are thought to be grounded in human physiology and cognitive

²⁶ Crucially, Hayes' (1996) constraints are not a direct reflection of the phonetic difficulty maps. The constraints are also marked by formal simplicity ('symmetry') and by their categorical nature, in contrast to the gradualness of phonetics.

²⁷ Alternatively, if it were shown that voiced geminates are not attempted by children in babbling, this could be one of the ungrounded constraints, which can be posited to exist alongside grounded constraints, and which are not based on experience with articulatory difficulty or perceptual confusability, but on the characteristics of the input data, i.e. on "systematic, consistent, long-term absence of a particular structure in the input data" Hayes (1996: 26).

capabilities, reflect the aspects of the environment in which linguistic constituents replicate. They can be thought of as reflecting the limitations on the replication process inherent in the external environment, that is in the characteristics of the speakers. But the environment in which linguistic constituents replicate is also made up of, more immediately, other linguistic constituents. Also, the copies of themselves which constituents spawn become the features of the environment in which further copying takes place. This can be captured by the other set of constraints which form part of the OT framework. Specifically, the relationship between the representations and their realizations is circumscribed by faithfulness constraints, which specify how much the output can differ from the input, and thus keep in check the discrepancies between mutants and their predecessors. Too big a discrepancy requires too much processing, and so faithfulness constraints exert a force to adjust the input to match the new output and thus drive change. There are numerous forces influencing the process of replication of linguistic constituents, and OT comes complete with a conceptual framework of formalizations which can be used to make the complex relationships between those forces visible.

From such a conceptualization, a view emerges under which phonemes, at least some markedness constraints and constraint rankings are replicators. Markedness constraints grounded in human physiology have as their environment both other linguistic constituents and the human body. They are then, most directly exposed to the biological aspects of humans as organisms. For the replication of phonemes, the environment constitutes of constraints and their rankings. It is only indirectly, i.e. through markedness constraints grounded in human physiology, that the human body is an environment for the replication of phonemes.

2.5.2 Further advantages to OT

It has been suggested that the evolutionary approach can profit from the formalization of OT. Likewise, the OT formalization of language change can profit from being placed within the evolutionary linguistic perspective. Apart from the modifications to OT presented above, which are welcome from the evolutionary perspective, two further ways in which OT can profit from being embedded in evolutionary linguistics can be mentioned.

First, on the conceptual level, phonology gains a clear ontological status when it is thought of as being encoded in brains of actual human beings. Granted, OT is a framework of formal phonology, and Kager (1999: 26), accordingly states that “[e]xplaining the actual processing of linguistic knowledge by the human mind is not the goal of the formal theory of grammar, but that of linguistic disciplines (such as psycholinguistics, neurolinguistics, and computational linguistics)”. However, two facts speak for bringing OT closer to the reality of what happens in the brains of speakers/hearers. The first is to be found within the field of OT itself. The very same handbook in which Kager voices the opinion on the irrelevance of the psychological details of language processing contains a chapter (1999: 296) on the learnability of OT grammars, so the assumption that there is a link between the formal model of competence and actual physically existing hosts of this competence is tacitly made. Admittedly, a formal theory of learnability is not the same thing as an empirically founded theory of phonological acquisition. Research in learnability seeks to formulate formal proofs to test particular hypotheses about Universal Grammar (McCarthy 2002: 202), whereas research into acquisition is primarily interested with how acquisition proceeds in actual speakers/hearers. But the goal behind the former is not irrelevant to that of the latter, and the assumption behind testing learnability has to remain that languages, in the end, must be learnable by actual human beings. The very existence of the line of research within OT which deals with the issues of learnability of grammar is testament to the wish to bring OT closer to the psychological reality, as the idea behind such work is that a model of learnable grammar is a better model. The second reflects the conviction of the author that a formal theory, while absolutely necessary to making explicit falsifiable claims, remains relatively uninteresting as long as it shies away from defining the mode of existence of the object under study in a way which would make it part of the physical world.

Second, modeling historical change as a relative change in frequency distributions of constituents allows for the presence of inter- and intra-speaker variability. A change modeled in OT gives the appearance of having swept through all the speakers at the same time, and thus excludes the interaction of older, unchanged competences, and the innovative competences. Such interaction, however, is not only possible, but even expected under the evolutionary view. Since in OT language change boils down to a change in the ranking of constraints, then representing a certain development takes the shape of formulating

successive constraint rankings, together with presenting motivations for the change from one ranking to the next. What results is an unbroken, continuous chain of rankings linking any two stages in a language. This may not be problematic from the point of view of OT if one is not interested in the implementation of the change but only in two idealized states of a language. However, if a change is to be placed in the communities of actual speakers, and this has to be the case if one wants to give a meaningful account of a historical change, then the issue of abruptness versus gradualness has to be addressed. The population perspective proposed here stands firmly on the side of the gradualness of change, and so is compatible with variationist sociolinguistics, which has found that change is diffused through communities, as argued by Labov (2001) or Trudgill and Chambers (1998). Change is known to spread from speaker to speaker, depending on the social structure of the actual groups these speakers belong to, and does not affect all speakers at once. What this means is that older, unaffected forms of a language variety can and do affect the later, innovative forms of this variety. Relevant for this second argument is the contribution by Dinkin (2012), who suggests that thinking of vowel shifts as phenomena occurring in the social space dissolves some of the thorny issues in this field of research, such as the question of the uniformity of shifts. Taking up Preston's (2008) finding that communities which a shift has diffused to show a more symmetrical pattern than the communities in which the shift originated, Dinkin proposes a view of chain shifting under which the development of a chain shift can be roughly divided into two stages. In stage one, a shift originates and is passed down through first language acquisition in the source community as a fully coherent development, with phonemes changing together. Then, if it is diffused to other communities, whose members acquire it later on in life, the pattern resulting from the shift becomes simplified and phonetically symmetrical, which may obscure the functional relationships between the original phonemes. Consequently, he argues that the (lack of) coherence of a pattern can be a matter of perspective, that is of which stage of a shift is looked at and which community.

3 THE RELEVANT DEVELOPMENTS AND A TENTATIVE ACCOUNT

As mentioned above, the starting point of the current investigation is the informal observation that starting with the Great Vowel Shift, which took place in the EModE period, the English vowel system seems to be in a constant state of flux, with many contemporary dialects of English likewise being characterized by vowel shifts. Before the verity of this observation is tested and a possible explanation is offered, Section 3.1 provides an overview of historical and contemporary vowel chain shifts in English, which are the focus of the hypotheses put forward in the present thesis. Section 3.2 reports on a metatheoretical study which investigated the validity of the assumption that the prevalence of vowel shifting has increased. Section 3.3 discusses the link between the evolution of rhythm and segmental developments, and leads up so Section 3.4, which maps out the core of the argumentation of the account of the onset of vowel shifting proposed in this thesis. The account presented in this last section of Chapter 3 involves a discussion of the usefulness of employing the concept of ‘exaptation’ in this account specifically, and in linguistics in general.

3.1 AN OVERVIEW OF VOWEL CHAIN SHIFTS IN ENGLISH

The onset of the increased tendency to undergo vowel shifts seems to have started with the Great Vowel Shift, a major reorganization of English vowels between the fifteenth and seventeenth centuries (Lass 1999: 80),²⁸ which is responsible for many discrepancies between Medieval and Modern English vowels. The following sections, accordingly, begin with an outline of the Great Vowel Shift, including a treatment of the controversial issue of the unity of the change. What follows is a brief presentation of another historical vowel shift, the Short Vowel Shift, and of contemporary vowel shifts.

3.1.1 *The Great Vowel Shift*

The vowel system of Modern English is markedly different from that of Middle English, and the most salient discrepancies lie in the correspondences between ME long vowels and their

²⁸ with sporadic attestations of the raising of /o:/ already in the fourteenth century (Lass 1999: 79)

Modern English reflexes. It was observed first by Luick (1921-1940) and then by Jespersen (1909–49) that the individual changes which have to answer for those discrepancies are not haphazard but form a pattern. This pattern, represented schematically in a version of the well-known diagram in Figure 4 below, is now traditionally referred to as the Great Vowel Shift (henceforth GVS). It is a series of raisings and diphthongizations undergone by the long vowels of Middle English. All non-high long vowels were raised, and the two already high vowels diphthongized. Both Luick's and Jespersen's account rests on the assumption that the process was characterized by internal causality, that is on the belief that the movements of the individual vowels across the vowel space were interrelated. This internal coherence is thought to account for the overall neat pattern.

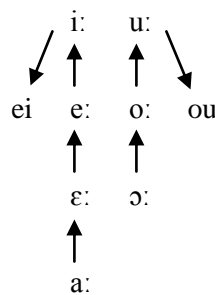


Figure 4: The Great Vowel Shift (after Lass 1999: 73)

This pattern was later obscured by the further lowering and centralization of the nuclei of the newly formed diphthongs as well as by a subsequent merger of the ME /e:/ and /ɛ:/, which, according to Lass (Lass 1992a: 73), lie outside of the GVS proper. A further difference between these values and those of PDE vowels is due to the nineteenth-century (Beal 2004: 137) diphthongizations of /e:/ and /o:/ to /ei/ and /ou/ respectively. Additionally, the first element of /ou/ was later centralized, resulting in its current RP value /əʊ/ (Gimson & Cruttenden 2008: 82).

Lass' claim that these later developments are not part of the GVS must be seen within the larger context of the debate on the unity of the GVS. Namely, the unity of the shift, that is the view that the pattern of the change is the consequence of movements of individual vowels exerting causal influence on the movements on other vowels, has been questioned. That it could be questioned is not surprising, since there is something peculiar about stipulating the

unity of a process which took about two hundred years to unfold. In particular, the functional explanation runs into trouble when accounting for the GVS, since in the span of those two hundred years, there must have been speakers who were able to communicate using phonological systems representing the intermediate stages of the GVS.

Stockwell and Minkova (1988b) list a number of unresolved controversial issues regarding the GVS namely “the inception problem” (what triggered the shift?), “the merger problem” (what is the importance of the preservation of contrasts for the shift?), “the order problem” (can individual stages of the shift be chronologically ordered, and if so, how?), “the dialect problem” (what was the importance of the developments in dialects outside of London for the understanding of the shift?) and go on to formulate their main thesis, which questions the established belief in the existence of the GVS, by laying out what they call “the structural coherence problem” (was the shift a unified phenomenon, a set of interdependent developments?). They claim that linguists’ attempts to address the inception and the order problem are rooted in their belief in the internal coherence of the shift, that is in causal relationship between individual changes. They cite Luick’s (1932, cited in Stockwell & Minkova 1988b) ‘displacement theory’, i.e. speakers’ motivation to maintain distance between potentially confusable phonemes as a mechanism which has been invoked to explain how individual phonemes can be seen to influence each other. They, however, refute its validity, since they do not believe that the preservation of contrast can play a role in sound change. Having come to the conclusion that the structural coherence problem cannot, or rather should not be overcome, then, they turn to more militant formulations of the issue, by saying that “[t]he Vowel Shift [is] the linguist’s creation through hindsight” and “a notable monument of scholarship that is in a real sense fraudulent” (1988b: 376). They see the remaining problems as pseudo-problems stemming from the, according to them misguided, belief in the unity of the shift. They propose an alternative analysis (1988b: 376) of the developments which have led to the modern reflexes of ME long vowels, suggesting that high vowels were diphthongal to begin with, but they were marked, and that dissimilation of the two elements of the diphthongs resulted in the modern diphthongs. The mid-high vowels were then raised due to markedness of vowel systems lacking high vowels, and the remaining high vowels merged with diphthongs they were closest to. Stockwell and Minkova (1988b) see what happened as a story of optimization of diphthongs (out of markedness

considerations), filling in of empty slots (in accordance with typological markedness) and mergers (which result not from structural considerations, but from the whims of “schoolmasters and language-conscious educators”). However, they claim that the individual developments cannot be seen as a unified, complex event.

One criticism which could be levied against their insistence on the independent nature of the developments of the high and mid-high vowels, is that their alternative also necessitates a kind of unity of the GVS. First, their own proposal rests on the assumption of related causes for the diphthongizations of the front and the back vowel. Second, the assumption that due to universal markedness considerations, a gap in the phonological system brought about by the diphthongization of the high vowels will necessarily involve a raising of a mid vowel is suspiciously close to a call for causality. In Stockwell and Minkova’s (1988b: 367) own words: “[w]e agree with Jespersen, and Chomsky – Halle, that diphthongization of high vowels took place first. Once that had occurred, then [...] it is clear that raising **must** [emphasis mine] occur for typological reasons, to restore a high vowel”. If the vacating of a slot in the vowel space must result in it being filled, then there is a sense in which vacating the slot causes the raising. Presumably, for Stockwell and Minkova (1988b) a unity could be assumed only if a movement of one vowel was the only reason for the movement of another vowel, without any independent causation, such as markedness, but would not be a convincing defense.

Lass (1992b: 150) argues for the unity of the shift, and his defense rests on two pillars. The first one is narrowing down the scope of the GVS to the shift involving four vowels only, ME long high and mid vowels. A coherence of the pattern excluding the later lowering of the nuclei of the new diphthongs as well as the *meat/meet* merger and the fate of the low vowel is much more defensible. The second line of defense one is the dialect-based observation, dating back to Luick (1896, referred to in Lass 1992b: 150) that northern dialects which lacked the back mid vowel due to fronting also lack the diphthongization of the high back vowel. Therefore, he concludes, the raising of the mid back vowel and the diphthongization of the high back vowel is best seen as forming a unit, and, by analogy, the same is true of the front mid and high vowels. He, however, though arguing strongly for unity, does not want to stipulate ‘causation’, but instead argues for “a particular kind of chain shift, in which segments in a given phonological subspace play musical chairs and don’t go

anywhere” (Lass 1988: 397), and he states that “[t]he term GVS denotes [...] that particular no-collapse shift that ends up with the Middle English long monophthong system intact, if phonetically displaced [...]” (Lass 1988: 397).

Stockwell and Minkova (1988a: 412) refer to this insistence on ‘unity’ but not on ‘causation’ “disingenuous” and, it seems, rightly so, since if the raisings did not cause the diphthongizations, then it is unclear what exactly this unity that the dialect data supports is supposed to consist in. According to Lass, the raising of mid vowels was a “trigger” (Lass 1992b: 152) for the GVS, but not its cause, and it is difficult to see what this distinction should mean.

This debate does not seem to have a clear outcome, even narrowing the GVS to the four vowels only, since neither account provides a definitive answer to the question of causation. Lass avoids this issue by referring to ‘unity’ instead, although his account does rest on causation after all, and Stockwell and Minkova argue feverishly against causation, even though, under their account, the diphthongization of the high vowels was a pre-requisite for the raisings of the mid vowels. Agreeing with Lass’ view that the GVS was a coherent development, the present analysis assumes further that it also involved causally related events. Further, the preservation of contrast is assumed to have played an important force in this development.

At any rate, the problem of coherence of the change appears in a new light when the population perspective is adopted, where additionally the population of linguistic constituents, and not the population of speakers, is the locus of change. While a conspiracy between generations of speakers separated by hundreds of years which would give a shift a certain direction is untenable, a change viewed from the perspective of the linguistic constituents which has a certain direction is absolutely conceivable. After all, the environment in which the vowels replicate (here: the effect of fixed lexical stress, rhythm) exerts its pressure constantly, regardless of individual (generations of) speakers. In addition to the long term effects of stress, in a population in which /e:/s surface as raised, i.e. as /i:/-like, those /i:/s will replicate well which are no more /i:/-like themselves, and such an

influence of a population of /e:/s on a population of /i:/s can also be accommodated in an evolutionary scenario without necessitating a generation-spanning conspiracy of speakers.

All approaches to the GVS mentioned so far rely heavily, if not exclusively, on system-internal factors for the explanation of the vowel shift. In contrast to these, there have also been attempts (Perkins 1977; Leith 2002; Lerer 2007) to account for the GVS by placing the weight of the explanation on the sociolinguistic context in which it took place. The fact that all these accounts refer to is that after French lost its status in England, it could no longer be used by the upper classes to distance themselves from the lower classes. As a result, social stratification came to be reflected in the emerging social dialects, as opposed to social languages. For Perkins (1977), it is ‘adaptive rules’, that is the modifications introduced by the speakers of the lower classes to their speech in order to imitate the upper classes, and their overgeneralization that initiated GVS. For Leith (2002), the London bourgeoisie, in order to distance themselves from the speech of the lower class, marked by features of Kent and Essex origins, exaggerated the fronting of the long /a:/ typical of Kent and Essex, a suggestion which is clearly preposterous. For Lerer (2007), the GVS is the result of particular dialectal variants as socially desirable, in the context of the variety of dialects in London at the time, and the emergence of a socially desirable standard. The scenario given by Perkins (1977) is the most worked out of these, and it seems to provide an overall plausible account. However, its reliance on the assumption that the GVS was initiated by the diphthongization of the long high vowels, as well as the early time of the shift undermine its plausibility. Leith’s (2002) account rests on assumptions which run counter to most philological as well as sociolinguistic research. Lerer’s (2007) proposition, for all the vivid presentation of the social environment of the time, does not propose any concrete mechanism. Its general outline would have to be that those dialectal pronunciations ended up selected to be part of the emerging standard which happened to result in the reorganization of the long vowel system. This huge linguistic reorganization would be entirely coincidental, and is thus nothing more than a strange curiosity. A serious problem with accounts which seek to replace internally driven change with external, i.e. social motivations is that, as pointed out by Lass (1988: 406), the various dialectal forms do have to come from somewhere. A detailed dialectal analysis attesting to the existence of the relevant dialect forms, as well as to their social ascension would be needed to flesh out the scenario. But even with that in place, the different

forms have to come into being due to endogenous factors, since all variant pronunciations cannot be posited to have existed from the dawn of time.

For all the deficiencies of the accounts reported on above, their common core, namely the emerging London standard simultaneous with the end of the role of French as the standard must have indeed provided a fertile breeding ground for linguistic change. A social dynamic in which a prestige form is first formed and then adopted by numerous speakers, is conducive to language change. Naturally, sounds are used not only to realize lexical contrasts, but also to signal social group affiliation and distance. This has relevance both for the vowel shifts as well as vowel mergers that took place around the GVS and that are the general concern of this thesis. Following the pioneering work of Labov (1966 [2006]) in the US and Trudgill (1974) in the UK, numerous studies have shown that differences in the affiliation to social groups correlate with differences in the realization of linguistic variables. And so, Bucholtz (1999) reports the resistance of California nerd girls²⁹ to the fronting of /u:/, as well as their use of fully released /t/s word finally. Eckert (1988) found a difference in the probability of use of the backed and lowered /ʌ/ between members of two social categories. Munson (2007) confirmed that listeners detect differences between the speech of gay and heterosexual men, as well as between the speech of lesbian/bisexual and heterosexual women. African Americans do not participate in the Southern Shift or in the Northern Cities Shift (Preston 2000). These findings suggest phonology is used in the enactment of social affiliation and distance. Since such sociophonetic variation is pervasive, it is bound to have an impact on the way in which sound change proceeds.

In the context of EModE sound changes, for instance, a group that has played a role in shaping the post-GVS mergers was that of ‘the Mopseys’. This name is an anglicized version of the Latin term *Mopsae* coined by Gil to disparage “a type of affected, over-delicate, hypercorrecting female speaker” (Lass 1999: 92). Their speech around 1550 was marked by a merger of ME diphthong, as in *days*, with the /ɛ:/, as in *seas*. In Popular London speech, however, ME /ai/ merged with /a:/, as in *daze*, rather than with /ɛ:/. Present-day English follows the Popular London pattern, so that nowadays *days* and *daze* rhyme, but *days* and

²⁹ i.e. female high-school students characterized by ‘nerd identity’, which is, according to Bucholtz (1999: 204) “a purposefully chosen alternative to mainstream gender identities which is achieved and maintained through language and other social practices”.

seas do not. The disdained pronunciation did not make it, then, and the speakers who shared Gil's disdain distanced themselves from the Mopseys by not adopting their speech pattern. Thus, the replicative success of various sounds was dependent not only on the system-internal co-replicators, but also on the social speakers that their hosts happened to be part of. In this case, the lineage embodied in the Mopseys did not succeed due to the stigma attached to them. It is in this way that sociolinguistic factors enter the scene, as an important environmental factor, which shares its importance with other, traditional 'system-internal' factors, rather than as the only set of factors to be considered important in sound change, as is suggested by Leith (2002) or Lerer (2007).

3.1.2 The Short Vowel Shift

Next to the Great Vowel Shift, another Early Modern English vowel chain shift was postulated by Schendl and Ritt (2002). Generally speaking, the Short Vowel Shift lowered and centralized EModE short vowels, and it forms a coherent pattern, which can be represented graphically as the one in Figure 5. There was one exception to the overall lowering and centralization; while /i, e, o, u/ lowered and centralized, the already low and central /a/ raised and fronted, since it could not be lowered or centralized any further. Interestingly, the exceptionality of /a/ in Short Vowel Shift parallels the exceptionality of the already high vowels /i:, u:/ in Great Vowel Shift, which could not be raised any further, and so diphthongized. Taking a "conventionalist and constructivist" point of view on the issue of evaluating historiographic accounts, in line with Lass' (1997) views on the tasks of historical linguists, Schendl and Ritt (2002) maintain that an account of a historical event should not be judged with respect to its truth value, as such a judgment is impossible to make, but on account of its fruitfulness. And so acceptance of Great Vowel Shift must lead to the acceptance of Short Vowel Shift, since if the reasons for the acceptance of Great Vowel Shift (presented below), be it given explicitly or accepted implicitly by the linguistic community, hold equally well when applied to the set of vocalic changes postulated to constitute Short Vowel Shift.

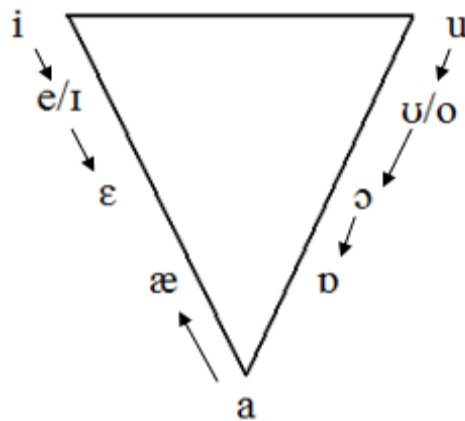


Figure 5: The Short Vowel Shift (after Schendl & Ritt 2002: 418)

Schendl and Ritt (2002: 411), by comparing the three vowel shifts proposed by Luick (1921-1940), one of which has come to be known as Great Vowel Shift, and two further, which have fallen into oblivion, identify four characteristics of GVS which distinguish it as a ‘good’ chain shift. First, it applied to an easily specifiable group of vowels, namely all long vowels of Early Modern English. Second, the individual changes affecting those vowels can be unified with a general statement. For Great Vowel Shift, this generalization is that all of the vowels involved raised, with the caveat that the already high vowels, which could not raise any longer, diphthongized. Third, a graphic representation of the developments yields a pattern. Fourth and last, there can be postulated, if only illusorily, a causality with each of the changes causing the next one.

3.1.3 Contemporary vowel shifts

It could be argued that the discovery of the shifts presented in the following, and the research that ensued, can be seen in a way as a gauge of the fruitfulness (or ‘conciliences’) of the Great Vowel Shift. Local varieties of English spoken in the following regions have been postulated to undergo vowel chain shifts: the Inland North of the United States (Labov & Yaeger 1972), New Zealand (Bauer 1979; Gordon, Hay & Mclagan 2008: 41), Southern United States – where two shifts are under way, namely the Southern Shift (Labov 1994 [2010]) and the Back Upglide shift (Labov, Ash & Boberg 2006: 127), Canada (Clarke, Elms & Youssef 1995), Australia (Cox 1999), South-East England (Torgersen & Kerswill 2004), Pittsburgh (Labov, Ash & Boberg 2006: 271) and Northern California

(<http://www.stanford.edu/~eckert/vowels.html>, Eckert & Mendoza-Denton 2006) among others. That the discovery of vowel chain shifts in these varieties owes its existence to the presence of Great Vowel Shift in the linguistic tradition is difficult to question. Indeed, the title of Laurie Bauer's article – "The second Great Vowel Shift?" – leaves no doubt as to the role of Great Vowel Shift in his investigation of the vocalic changes underway in New Zealand English. For others, the existence of the Great Vowel Shift as a recognized entity also arguably played a role. In a wider perspective, the fascination with vowel chain shifts has contributed to growing sophistication of linguistic methodology and to highlighting the links between language and society, with both aspects present in the work of William Labov.

At any rate, the picture emerges that English vowels have been implicated in a considerable number of chain shifts. This started in the Early Modern period and is still happening. Given this remarkable difference between pre-EModE and post-EModE states of affairs, one is tempted to ask whether English changed at that time in some fundamental way.

3.2 STUDY 1: ENGLISH HAS BECOME A VOWEL SHIFTING LANGUAGE

As is clear from the discussion above, English is at present undergoing numerous vowel chain shifts. A perusal of vocalic changes which took place in earlier stages of the language suggest that it was not the case, or at least not to the same extent, in the past. This, still rather impressionistic, claim is investigated more rigorously in the following, by means of a metatheoretical study based on extant accounts of vocalic changes in English.

The empirical claim that forms the base of the hypotheses to follow is that English has become more likely to undergo vowel chain shifts in the course of its history. The most straightforward way to verify this assumption, then, would seem to be to measure the increase in the number of chain shifts over time. However, a set of qualitative changes the size of the Great Vowel Shift, is too big an entity to measure the rise in its prevalence over the course of the history of English. To surmount this difficulty, another unit of measurement has to be considered. Consequently, for the sake of the following investigation, a method of tracking the rate at which vowel chain shifts may befall a language has been devised which has as its basic units changes undergone by individual vowel phonemes. To this end, both

qualitative and quantitative changes undergone by individual vowel phonemes over the course of the history of English were charted. Of all the possible changes that a vowel sound might undergo, it is unconditioned changes not resulting in merger (from now on also referred to as Type 1 changes) that might constitute the building blocks of a vowel chain shift. Now, these changes not resulting in merger are the sorts of changes which observers might, given there is enough such changes, construct as forming a pattern. This pattern might then be interpreted as a vowel chain shift. Without even considering the question of causality or internal coherence, when a number of these changes take place, it must by necessity form something approaching a pattern which is suggestive of internal coherence. Since what is defined post facto as a vowel shift is a series of events which do not result in mergers, then the building blocks of chain shifts must be Type 1 changes. The obverse must not follow logically, that is an increase in Type 1 could be conceived of as a necessary but not a sufficient prerequisite of vowel chain shifts. Taking the limitations of the vowel space into consideration, however, it is plausible to assume that when a number of phonemes are moving around that space, a pattern will emerge.

Changes resulting in mergers are, by definition, not legitimate candidates to be part of a chain shift. Conditioned changes, which involve a change in a clearly defined phonological environment, leave at least some lexical items containing the vowel phoneme in question (for which the phonological environment cannot be clearly identified) unchanged, and so are no candidates to be part of a chain shift either. Following this assumption, the question of whether English has become a vowel shifting language has been re-phrased as the question of whether English has come to undergo more unconditioned changes not leading to mergers. A catalogue of known vocalic changes, divided into (a) unconditioned changes not resulting in merger and (b) other vocalic changes, has been compiled on the basis of textbook descriptions and critical investigations. A statistical analysis based on these data has indeed confirmed that the number of such changes has significantly increased.

3.2.1.1 Charting vocalic changes

All vocalic changes defined above were gleaned from the following sources: For the period from the very beginning up to the eighteenth century, CHEL (Cambridge History of the English Language) was the main source, or its chapters pertaining to phonology, namely

(Hogg 1992a), (Lass 1992a) and (Lass 1999). Additionally, for the OE period (Hogg 1992b) was used to clear up some timing issues. For the later stages, Barber (2006) was used for Early Modern English, Beal (2004) for Late Modern English, and Cruttenden (2008) for Present Day English.

For the present study, only one lineage of English, or the closest one might get to it based on historical evidence, was charted. South-Eastern British English was chosen for the present state of English, and the history of its ancestral varieties was traced. This variety was chosen due to the abundance of documentation, both for the present-day stage as for the historical stages. The assumption was that only one lineage should be investigated, since the differing number of varieties documented for different periods would inevitably skew the number of changes to be measured. Most clearly, the number of varieties documented for Present-Day English would automatically have to lead to an increase in Type 1 changes over time. By trying to trace a single lineage of varieties, this effect was controlled for.

	16th	17th	18th
	I	I	I
	i	i	i
	i:	i:	i:
	ε	ε	ε
	ε:	ε:	ε:
	e	e	e
	e:	e:	e:
No. of vowel phonemes	4	5	4
No. of Type 1 changes	2	1	0
No. of other changes	0	1	0

Table 1: Portion of the table of vocalic changes. Vowel phonemes present in the vocalic system in a given century appear in black, those absent from it appear in grey. Type 1 changes are indicated with thick arrows, other changes are indicated with thin arrows, a continuation of a phoneme into the next century is indicated with a line.

To enable quantification, the period from the earliest reconstructed OE changes (fifth century) to the present day (twentieth century) was divided into centuries. This means that each vowel change was charted as a transition from one century to the next, where the latter was the time when the change was well-established. The number of vowel changes was then counted for each century. For example, for the particular slice of the chart in Table 1, two Type 1 changes have taken place on the way from the sixteenth to the seventeenth century. The number of vowel changes was ascribed to the century when the change originated (here: the sixteenth century), and not when it was established (here: the seventeenth century), as a reflection of the belief that the structural properties of the vowel system are among the environmental factors influencing each other's replication, and so the properties of the vowel system in a given century may influence the developments from that century to the next. This conviction stems from the spirit of the evolutionary perspective, under which vowels do not replicate in a vacuum but in an environment in which other vowels also replicate, and there replication is believed to be influenced by the composition of the whole vocalic system. However, allocating the number of changes to the century in which the effects of a shift are already manifest (here: the seventeenth) as opposed to the one in which they originated (here: sixteenth) would not have affected the results of the study. It would move the numbers of shifts wholesale one step to the right, so it would not affect the rate of the increase of Type 1

changes over all changes. What is essential is that this marking is done consistently, and not its particular positioning according to one convention or the other.

The changes were divided into two types: the above mentioned unconditioned changes not resulting in merger, e.g. <eo> > /ø/, 11th century; /o:/ > /u:/, 16th century; and other vocalic changes, including unconditioned changes resulting in merger, e.g. /ø/ > /e/, 12th century; conditioned qualitative changes e.g. /a/ > /au/ before tautosyllabic /l/, 15th century; and conditioned quantitative changes e.g. /a/ > /a:/ in open syllables, 13th century. The entire chart is presented in Appendix 2. Table 1 above presents a snippet of it to illustrate the principle. From the method of ascribing the changes to the centuries in which they are assumed to have originated it follows that for the last column charting a particular period the number of changes will by necessity be zero. That is why the numbers for the last century in the entire table, that is for the twentieth century, are also zero.

3.2.1.2 Analysis

To measure whether the number of Type 1 changes has increased over time, their proportion to all other changes for each century was calculated. This was done to make sure that a rise in their frequency does not simply reflect a potential rise in the overall rate of language change. It is possible that languages undergo more change at some stages in their development than at others, and measuring one particular type of change against all other possible changes was meant as a safeguard against this influence. The number of Type 1 changes originating in each century next to the number of all changes originating in each century, as well as the proportion of the former to the latter is presented in Table 2 below.

Century	Type 1 changes	All changes	Type 1 changes / all changes
5	0	7	0
6	0	13	0
7	0	3	0
8	0	2	0
9	0	18	0
10	2	4	0,5
11	1	5	0,2
12	1	18	0,06
13	0	1	0
14	0	1	0
15	7	11	0,64
16	10	17	0,59
17	2	8	0,25
18	4	9	0,44
19	7	9	0,78

Table 2: Number of Type 1 changes, of all changes, and the proportion of Type 1 changes to all vocalic changes over time

To enable better visualization of these results, Figure 6 below presents the values of the same proportion over time as a scatter diagram. Pearson's product-moment correlation test yields a significant correlation with $p = 0.003$ and $cor = 0.7$, which corresponds to a large effect size.

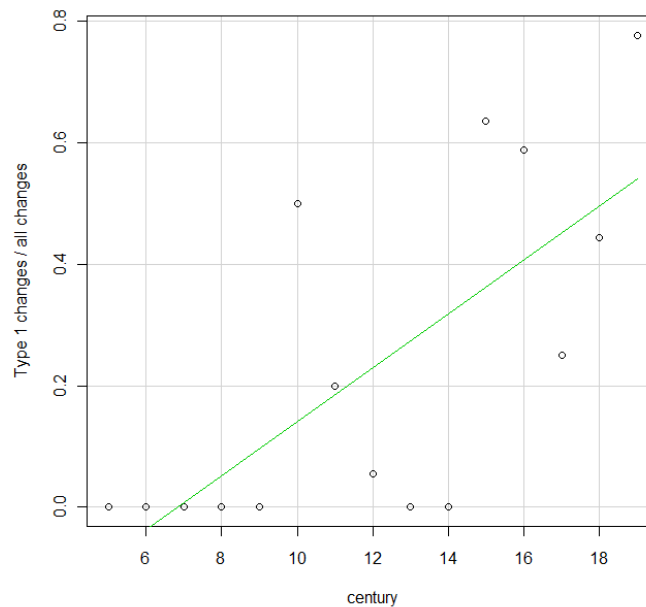


Figure 6: Proportion of Type 1 changes to all vocalic changes over time

These results, i.e. the finding that the number of unconditioned vowel changes not resulting in merger has increased, are consistent with the assumption that English has become a vowel-shifting language.

In anticipation of a potential summary explanation which ascribes a rise in the number of vowel chain shifts to an increase in the number of vowel phonemes, a look at the changing number of vowel phonemes in English is in place now. Figure 7 below presents the number of vowel phonemes for each century. With so few data points, and therefore low power of a potential statistical test, it is impossible to definitively attest to a lack of an increase. However, there is nothing about their distribution over time that would point to the legitimacy of postulating such an increase in the first place. What is more, as the large number of (qualitatively distinct) vowel phonemes can be posited to be characteristic of stress-timed languages (cf. Donegan 1993), and in the case of English, the proliferation of vowel qualities was already the first step of the GVS, invoking the large number of vowel phonemes is not really a counterargument to the mechanism proposed here.

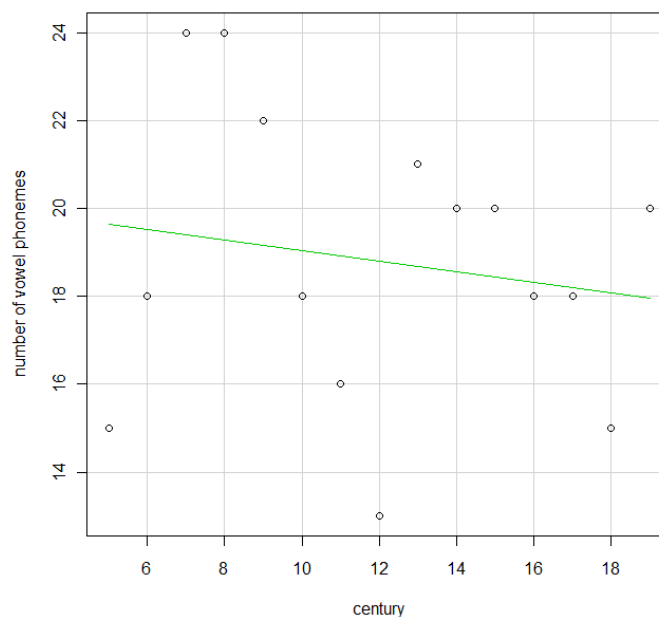


Figure 7: Number of vowel phonemes in English over time

3.3 THE CO-EVOLUTION OF RHYTHM AND VOWELS

It is proposed in this thesis that the GVS is closely linked to a set of events which together have resulted in the fuller expression of trochaic rhythm in English, and thus in making English a more stress-based language. First of those is the reduction and loss of vowels in unstressed position (Section 3.3.1). Now, the traditional dichotomy of stress-timing and syllable-timing has been repeatedly put into question (as discussed in detail in Section 3.3.2 below). Something of a consensus is that all languages are stress-based, and that this property is expressed in speech production or perception to varying degrees in different languages or at different stages in the development of one language. Consequently, it is not claimed here that English radically changed its rhythmic type at a given point in its history. An attempt to do so was undertaken by Markus (1994), who sees ME as syllable-timed and places the transition to ME stress-timing in EModE. Ritt (2012), in contrast, sees already ME utterance rhythm as isochronous with regard to stresses, and the evidence of clash and lapse avoidance in ME supports this stance. However, since there is no strict dichotomy of stress- and syllable-timing, a point of transition between the rhythmic types cannot and need not be identified. Instead, a move towards a more robust expression of stress-based rhythm is discernible in the history of English, as most³⁰ successive changes forming part of the ‘Great Trochaic Conspiracy’ (2004: 289) made more English morphemes be expressed as better trochees. Previous attempts to explain the pervasiveness of vowel shifting in English as a typological characteristic are presented in Section 3.3.3. Section 3.4. introduces a holistic perspective on the quantity adjustments as well as their link to vowel shifting. A first, tentative scenario of these developments, seen as a case of ‘exaptation’ is then proposed.

3.3.1 The reduction and loss of unstressed vowels

The mechanism behind the replacement of quantitative differences by qualitative differences which, as argued in this thesis, brought about the various chain shifts described above is closely linked to another development in English phonology, namely the reduction and loss of unstressed vowels. As both changes, namely qualitative shifts and vowel reduction and loss, are linked to lexical stress (and as both developments have been influenced by their co-

³⁰ With the notable exception of final schwa loss, cf. Section 3.3.1, as well as Section 4.3.

evolution with the trochee) further discussion requires at least a short presentation of the basics of the history of final schwas.

Minkova's (1991) monograph *The history of final vowels in English* provides a wealth of information regarding the fate of schwas in final syllables and in absolute final position. The survey of important accounts of schwa loss that she provides converges on the following timing: the beginning of the change in the North between 1100 and 1250; the spread of it through Midlands and the South between 1250 and 1350, and a complete loss, including the South, between 1350 and 1400. The same dating is taken over by Lass (1992a: 79). But schwa loss was preceded by a number of changes weakening vowels in unstressed position.

First, quantitative oppositions in unstressed syllables, present in Germanic had been leveled out by earliest OE times, and only short vowels could be found in unstressed position in OE (as evidenced e.g. by Gothic *a*-stem nom. pl. *-ōs* vs. OE *-as*) (Lass 1992a: 77). In Late Old English, i.e. in tenth century, only three vowel qualities were to be found in unstressed syllables, namely /e/, /o/ and /a/, with the other vowels present here before merging with one of these three qualities³¹ (Minkova 1991: 89). By the eleventh century, the three qualities (excluding /i/) merged into one, which was "usually spelled <e>" (Lass 1992a: 77). Though the exact phonetic quality of the vowel resulting from the mergers is not easy to determine, it is fairly clear that vowels in unstressed position were non-distinct.

Although the exact timing of the sequence of events presented in this paragraph, as well as of schwa loss, is notoriously difficult to determine, there seems to be a consensus that the beginning of schwa loss roughly coincides with the loss of distinctiveness of vowels in unstressed position. This historical proximity, and, possibly, causal interdependence, forever linked the discussions of vowel reduction with discussions of schwa loss.

Most of the accounts of schwa loss Minkova (1991) cites, notably the earliest and very influential accounts of Morsbach (1896) and Luick (1921-1940), invoke the root-initial stress as the cause of schwa loss. The basic idea is that strong root initial stress corresponds to weak stress on the following syllables, which results in lenition. Even Horn and Lehnert (1954),

³¹ With the exception of derivational morphemes such as *-ig*, *-ic*, *-isc*, *-ing*, *-iht*, and *-lic*, in which /i/ did not merge with /e/ (Minkova 1991: 89).

who were the first to recognize the role of morphology in schwa loss, saw root initial stress as a condition for the loss of the vowel in the following syllable. According to Minkova (1991: 28), such phonetically-based accounts miss the very important distinction between reduction and loss. She observes that while the influence of stress on reduction cannot be denied, invoking stress does not suffice to explain loss. She identifies two questions which show this weakness. First, it remains to be shown why the stress was increasingly strong. Second, other Germanic languages, likewise having root-initial stress, have not undergone schwa loss. Hence, a full account cannot be limited to the phonetic motivation, but needs to involve other factors, such as “morphological insufficiency, analogy, prosodic organization, external influence [...] the syntactic patterns in Middle English” (Minkova 1991: 30).

The interplay with morphology, in general terms, is that the onset of schwa loss was enabled in nouns and verbs by the high level of syncretism, i.e. low level of morphological distinctiveness, caused by the mergers of weak vowels in one quality (Minkova 1991: ix). This low functionality, in turn, was brought about by the mergers, that is by (phonetically-driven) phonology. Thus, developments in phonology influenced morphology, and morphology later fed back to phonology, enabling schwa deletion, as there was no pressure to retain schwas, which were not particularly useful from the point of morphology anyway. The development of schwa loss itself, however, can be seen as largely prosodic in its causation. The earliest cases of schwa loss show that it was confined neither to a particular segmental nor morphological environment (Minkova 1991: 155). And the cases of late schwa preservation in weak adjectival inflections were not due to the morphological functionality of schwas, but due to metric reasons, which is made clear, for instance, by high rates of disyllabic forms among weak adjectives, which resulted from schwa preservation in monosyllabic roots and schwa deletion from disyllabic roots, regardless of grammatical motivation for the presence or absence of the morpheme (Minkova 1991: 171).

As schwa loss resulted in an increased number of monosyllables in English, it stood in the way of isochrony. A predominance of disyllables in the lexicon, with a strong first element and a weak second element, a feature of English in the OE period (Getty 2002, cited in Minkova 2006: 102), was conducive to trochaic rhythm. The proportion of monosyllables to disyllables rose through the ME period, with schwa loss dramatically contributing to the rise in the number of monosyllables. As a result of this change, the incidence of stress clashes

must have risen. These issues lie at the center of the account presented here, especially in Section 4.3.

3.3.2 Rhythm-based typology: on stress-timing and syllable-timing

Both the accounts of schwa loss recounted in the previous section and the claims that vowel shifting can be a typological characteristic of a certain language family refer to a rhythm-based typology. The feasibility of devising such a typology is discussed below, with the main outcome being that stress-based isochrony is, to some extent, a feature of all languages, a feature realized to varying degrees. The extent to which it is realized in a given language will fall on a scale, rather than in one of two polar opposites. Through patterns of historical changes, a language's position on this scale can shift, which is what is argued to have happened with English.

The idea of rhythmic classes goes back to Pike (1946), who observed that English can be differentiated from other languages, such as Spanish, on the basis of its rhythm, that is its durational properties. In English, he postulated, there are equal intervals between stresses or rhythmic feet, which results in a rhythm he dubbed stress-timed, whereas in languages such as Spanish, there are equal intervals between syllables, which results in a rhythm he dubbed syllable-timed. Later, Abercrombie (1965; 1967) extended Pike's dichotomy to claim that all languages of the world can be divided with regard to their rhythm, and that each of them will fall either in the stress-timed or the syllable-timed type. For him, this dichotomy was absolute and allows of no intermediate stages. This ties in with the proposed source of the distinction. He claimed that there was a physiological basis for the dichotomy, with either 'chest-pulses' or 'stress-pulses' recurring at isochronous intervals.

This very strong claim, making very clear verifiable predictions, attracted a lot of attention and has been repeatedly put to the test. Ladefoged (1967) refuted the physiological basis proposed by Abercrombie. Numerous experimental studies have failed to show that the intervals between stresses are equal in stress-timed languages, or that the intervals between syllables are of equal length in syllable-timed languages (e.g. Roach 1982; Dauer 1983, 1987).

These negative findings have led to a number of responses. One view is to see isochrony as a perceptual phenomenon (Lehiste 1977). Another is to view rhythm not as a phonological primitive, but as a property emergent from an interplay of a number of phonological traits, and as such, it is a continuous cline rather than a two-way distinction. One attempt to do so is Auer (1993). His investigation starts by deductively formulating two prototypes of word languages and syllable languages, with languages lying between those two prototypes being apparently in a transitional stage. He gives four illustrative examples of genetically related language pairs, where one of them has moved away from one end of the cline into the other, namely Italian and Portuguese, Uzbek and Turkish, Classical Mongolian and Khalkha and RP/Standard American English and West Indian Creoles. Yet another is to view isochrony as a tendency, realized to varying degrees in various languages (Beckman 1992; Laver 1994). This last approach seems to be very attractive and tractable from the OT perspective. As a first approximation, one could posit the existence of a constraint requiring the alternation of weak and strong units of timing, whose relative ranking would determine to what degree this alternation is actually manifested in output forms. Or, to anticipate two constraints used later on, one could argue that a constraint requiring that stressed syllables be heavy (STRESS-TO-WEIGHT), together with a constraint requiring that the leftmost syllable in a foot be the head syllable (RHTYPE=TROCHAIC), when ranked high, would result in a high incidence of forms realizing the trochaic foot. However, as argued in Section 4.2.3, positing the relative ranking of constraints yielding trochaic feet does not go all the way toward explaining a long-term tendency for more and more forms to realize the trochaic foot. A more general solution, from an evolutionary perspective, involves positing that constant articulatory and cognitive pressures seep into phonology, expanding the range of forms in which the trochaic foot is realized. New forms complying with the trochee are better able to replicate, as more and more forms expressing the trochee are in the environment already.

Attempts at a rhythm-based typology have not been abandoned faced with lack of experimental support from the studies which looked for phonetically present correlates of isochrony between phonological units (syllables and feet). It is not the case, though, that the continued currency that stress-timing and syllable-timing enjoy in the linguistic community is due only to tradition and the intuitive appeal that these notions have. First of all, the idea that rhythmic organization is crucial to language production as a complex motor activity is

not contested, and it seems to be in line with the findings of neurolinguistic studies into rhythmic organization (Sakai, Hikosaka & Nakamura 2004). Second, there are at least two sources of support for the hypothesis that languages differ in their rhythmic organization along a scale. These are studies in speech perception – by children, e.g. (Nazzi, Bertoncini & Mehler 1998), but also by tamarin monkeys (Ramus 2000); and phonetic studies into isochrony in the speech signal (e.g. Ramus, Nespor & Mehler 1999, Low, Grabe & Nolan 2000; Grabe & Low 2002).

One strong source of support for the reality of rhythmic classes is research on the perception of languages by children. Numerous studies show that newborns can distinguish between sentences of their mother tongue and those of another language, belonging to a different rhythmic class (for an overview, see: (Ramus, Nespor & Mehler 1999). In some of these, including Nazzi (1998), the sentences to be discriminated were low-pass filtered, to make sure that the infants rely on prosodic cues for the discrimination task. Nazzi et al.'s (Nazzi, Bertoncini & Mehler 1998) study provides particularly strong support for the hypothesis that newborns use prosodic cues to assign the ambient language to a particular rhythm-class. In their experiment, French newborns were able to distinguish between English and Japanese sentences, but not between English and Dutch sentences. It seems plausible, then, that infants are capable of assigning languages to which they are exposed to particular rhythm classes. The question that remains, however, is what kinds of cues they use to do so.

Taking this as their starting point Ramus et al. (1999) looked for a basis for rhythm classes in the speech signal itself. They measured three parameters, namely V%, that is the percentage of vocalic portions relative to utterance length, ΔC , that is variance in the duration of consonantal portions, and ΔV , that is variance of vocalic portions. Using these, they found support for the idea that languages can be grouped according to the degree to which successive vocalic and intervocalic intervals vary in duration. Low et al. (2000) improved on a weakness of that study – namely on the fact that it did not take the alternation between vocalic and consonantal portions into account – by developing PVI (Pairwise Variability Index). PVI measures the differences between durations of contiguous units, where a low value means that they are of exactly the same duration, and a high value that they are very different in duration. Applying the same metric to a number of languages, Grabe and Low (2002) conclude that their study supports “a weak categorical distinction between stress-

timing and syllable-timing”. These studies can be placed within the context of a wider group of recent investigations into the possibility of a rhythm-based typology (Wagner 2001; Cummins 2002; Torgersen & Szakay 2011). Despite their shortcomings as models of rhythm in general pointed out for instance by Gibbon (2004), they do seem to suggest the existence of rhythm-based classes.

More importantly for the present thesis, basically all the studies referred to so far have confirmed the, depending on the terminology employed, stress-timed (Ramus, Nespor & Mehler 1999) or stress-based (Dauer 1983) character of the rhythm of English. Even though a typology encompassing all, or even a group of, the world’s languages seems to be hard to come by, all investigations seem to be placing English comfortably sitting on one end of the spectrum, whatever the specifics of the particular typologies are.

Also, not without relevance to English, Schiering et al. (2012), though fairly critical of previous attempts at deriving a rhythm-based typology, note that some cross-linguistic generalizations do hold true, among them the ‘stress cline’.

[T]he following cross-linguistic generalization: the stronger stress is realized phonetically, the higher its impact on the phonological system. This means that for a language with strong phonetic stress, expectations are high that it will also show segmental effects of stress (vowel reduction in unstressed syllables, vowel lengthening in stressed syllables, and consonant changes triggered by the absence or presence of stress)

Additionally, Fear et al. (1995) have shown the reality of a binary, categorical strong syllable/weak syllable opposition in English in perception.

With regard to the approaches seeing rhythm as an emergent property, English has to be seen as becoming more and more stress-based, in light of developments such as the reduction of unstressed syllables, or decreasing phonotactic complexity.

Whether it is ultimately found that there are global rhythmic types, or more local correlations, the explanation of the mechanisms in which the development proceeds will have to refer to some concept or other resembling a long-term ‘conspiracy’. As mentioned earlier, cumulative, seemingly goal-oriented developments spanning numerous generations of speakers pose no threat whatsoever to an evolutionary account of language change. Here, it is assumed that the “weak categorical distinction between stress-timing and syllable-timing”

hypothesized by Grabe and Low (2002) results from the difference in the degree to which a general tendency for isochrony is realized in different languages, or different stages in the development of a single language.

3.3.3 Typological perspective

It has been noted in the past that some languages are more likely to undergo vowel shifts than others. Donegan (1985, reported in Awedyk & Hamans 1989) observed that vowel shifts occur in some languages more frequently than in others. They are frequent in Germanic, South-East Asian and Chinese languages, and infrequent in Finnic, Hungarian, Japanese, and some languages of India. She identified timing as the most important typological difference between shifting and non-shifting languages, with shifting languages being stress-timed and non-shifting languages being syllable- or mora-timed. The idea that the rhythmic organization of a language is linked to its likelihood to undergo vowel shifts, though without the necessity to assume a strict dichotomy between stress-timed and syllable-timed languages, is pursued in the account presented in the present thesis.

Donegan, having noticed a number of striking differences between two related language families, Munda and Mon-Khmer, pursues in a number of articles (Donegan & Stampe 1983; Donegan 1993; Donegan & Stampe 2004) the idea that these differences can be ascribed to the different rhythms of the two families. She identifies the following to be characteristic of stress-timed languages: the presence of large vowel inventories, and the historical tendency of their vowels to undergo diphthongizations, vowel reductions and vowel shifts. These differences are found not only between Munda and Mon-Khmer, but also other, better known languages fall into of the two categories, either stress-timed or syllable (or mora) timed. Donegan and Stampe (2004: 20) have refined the binary opposition of stress and syllable/mora timing somewhat, by stipulating that the tendencies for isochrony at all levels (mora, syllable and word) are there in all speakers at all times, but that “they are forces by the structure of their languages to yield on one or more of the principles”.

Diphthongization, a natural process which enhances the characteristics of a vowel, is given free rein in stress-timed languages due to the longer time devoted to the pronunciation of stressed vowels. The implication for the GVS would be, then, that the diphthongization of

the stress-timing EModE high vowels started the shift. The view on the ordering of the GVS adhered to here, however, is that the diphthongization of the high vowels did not precede the raising of the high-mid vowels. Since the earliest spellings of the raisings of /o:/ predate those of the diphthongizations of high vowels (Lass 1999: 79), and since in the dialects in which /o:/ fronted, rather than raised, /u:/ did not undergo diphthongization (Lass 1992b: 150), it is unlikely that the GVS started with the high vowels. Hence, this precise mechanism cannot be accepted. The general view that “[t]he rhythmic type of a language is a pervasive influence in its living phonology, and since phonological processes apply to rhythmic domains, they are strongly linked” (Donegan & Stampe 2004: 19), however, is adopted here.

The view that English vowels are likely to undergo shifts, and that there is a constant trigger for his ‘tendency’ goes back at least to Lotspeich (1921; 1927). According to him, the fixing of lexical stress in Germanic, together with the manner in which stress manifests in Germanic, make vowel shifts likely to recur. He suggests a physiologically-based scenario, where stress concentrates on the initial portion of a vowel, leaving all but the beginning of the vowel unstressed, which causes its raising. As the initial portion becomes smaller and smaller, the entire vowel takes on the value of the reduced second half, and so, changes like /e:/ > /ei/ > /i:/ and /o:/ > /ou/ > /u:/ take place. The same mechanism is supposed to hold for high vowels. The shrinking initial stressed portion of the vowel falls, because “the sudden initial stress-attack catches the tongue while on its way up to this high position” (Lotspeich 1921: 211). Though the observation that the fixing of lexical stress in Germanic and its ‘strength’ figure prominently in the advent of vowel shifting in Germanic is kept here, the mechanistic explanation of Lotspeich, which sees all the explanation in physiology cannot be accepted. For one thing, the peculiarities of the physiological mechanism it proposes are rather tentative, especially the supposition that the part of the vowel which is given most emphasis, that is the initial portion, reduces temporally. For another, Lotspeich’s account would predict only the occurrence of raisings and diphthongizations, in contrast to an account describing an advent of general instability of the vocalic system, which undergoes also changes like frontings and lowerings, which is attempted here.

The idea that the nature of stress in Germanic is linked to its likelihood to undergo particular kinds of changes is pursued by Iverson and Salmons (2003). Focusing on the

persistence of consonantal developments in Germanic, they postulate a change in the nature of stress in early Germanic, which they claim to have changed from a pitch accent to a duration/intensity based accent. They argue that a new ‘articulatory setting’, which permeates all aspects of phonology, underlies both a change from pitch accent to a duration/intensity based accent, as well as a change from a voiced/voiceless to aspirated/unaspirated contrast in stops. The latter change, that is a different realization of contrasts in stops, is supposed to account for the persistence of consonantal changes, as the unaspirated, or ‘passively voiced’ stops are ‘repeatedly enhanced’. This change in the nature of stress might also be responsible for the persistence of chain shifting in Germanic (though not necessarily along the lines proposed by the authors themselves, who posit implications for tense vowels only, in parallel to their proposal to the ‘passively voiced’ unaspirated stops).³² It can be argued that since one of the main correlates of stress came to be duration, the entrenchment of durational properties of stressed vowels came to interfere with inherent vowel duration due to length. Short stressed vowels, which would before that change be stressed by means of pitch, now came to be phonetically lengthened. Thus, a first step toward undermining the role of length as a contrastive feature in vowels was taken. In a later development of this trend, once pairs of vowels came to be distinguished phonetically by quality, and phonologically by [±tense], the vocalic system was set motion, as the previously qualitatively similar vowels started drifting apart. This later development is at the center of the discussion in Section 3.4, and of Chapter 4.

The view that some languages are more likely to undergo vowel shifts than others was also expressed by Stockwell (1978). He (1978: 345) maintains that English inherited the vowel system of Proto-Germanic, where no pure long vowels were present, but only simple vowels, ingliding and outgliding diphthongs. The presence of diphthongs in the system is thought to be responsible by the ‘perseverance’ of vowel shifting in English. This line of argumentation is continued by Minkova and Stockwell (2003: 187), who conclude that “[t]here are good phonetic reasons for the systematic development of diphthongs out of long vowels and vice versa within the vowel space. The resulting four changes: diphthongization, monophthongization, chain shifting, and merging, are distinct and differently motivated

³² Jacewicz et al. (2006) explicitly pursue the idea that prosodically induced strengthenings in English might account for the persistence of vowel shifting in English.

subtypes of the ‘umbrella’ type of phonological change known as ‘vowel shifts’”. However, in view of the results of Study 1, suggesting that the rate of vowel shifting has increased in English, it might prove fruitful to pursue the question of how this likelihood to undergo vowel shifts might change over time, rather than to assume it as a constant in the history of English, or even of Germanic.

Not without interest for the unified scenario proposed in the present thesis, Stockwell & Minkova try to place the GVS in the wider context of similar changes in the Germanic family. They define four “dynamic properties” which were behind the individual changes of the GVS, namely (1) alternation between long vowels and diphthongs (2) dissimilation of components of diphthongs (3) monophthongization once the diphthongs have become optimal (4) dialectal variation in the realization of vowels. These

define a certain type of vowel system within which such changes (**not** coherent shifts) are characteristic. These changes occur, recur, continue, and have gone on in this language family certainly since West Germanic times, possible since Proto-Germanic times, and most certainly are still going on in much the same way throughout at least the English-speaking world. (Stockwell & Minkova 1988b: 371)

Quite apart from the issue of the importance of the individual “dynamic properties”, the existence of a tendency would still require an explanation. This is a goal lying outside the scope of the present thesis. Instead, on a smaller scale, a closer look at the specific ways in which other features of English have influenced the development of its vowel system is addressed in the following section. The linking of the emergence of vowel shifting in English, or rather of speeding it up, to specific other developments in the language is seen as a worthwhile pursuit.

3.3.4 A unified perspective

Elements of the unified perspective, bringing together the long-term developments of vowel length with qualitative vowel changes, argued for here have been articulated before. The ‘holistic’ perspective of seeing the many quantitative changes as related was expressed, for instance, by Lass (1974), as recounted below, and a link between the erosion of vowel length and the onset of vowel shifting was postulated by Trnka (1982).

As for the interrelatedness of the various qualitative changes, Lass (1974) argues that a ‘length conspiracy’ has been operative in English. Taking as his starting point the idea of rule conspiracies posited for synchronic phonology by Kisseberth (1970), he extends it and claims that conspiracies can be seen to operate in historical processes as well. In short, he observes that a number of changes in the sound system of West Germanic and continuing in English, such as West Germanic Final Lengthening, Old English Quantity Adjustment, Pre-Cluster Lengthening, Early ME Quantity Adjustment and Open syllable Lengthening, have limited the number of contexts in which phonemic vowel length oppositions are maintained. In view of the fact that they seem to have similar ‘goals’, he posits that they have to be related, rather than coincidentally contributing to the end result. To account for this relatedness, he claims that these changes have been motivated by their ‘final cause’, i.e. teleologically, along the lines of the now discredited evolutionary thinker Teilhard de Chardin. De Chardin’s goal-directed version of evolution is a religiously motivated, “confused and ultimately self-contradictory alternative” to Darwinism (Dennett 1995: 320). Furthermore, it should be reminded that rule conspiracies were seen by phonologists as an embarrassing problem to be done away with rather than as an interesting finding to be recognized and built upon. As noted earlier, it is heralded as a great advantage of OT over derivational frameworks that it can dispense with such conspiracies by its exclusive use of output constraints. And while the choice of the optimal candidate can in a sense be seen as a process which is driven by its goal, that is a candidate is chosen so that it satisfies the particular constraint ranking, no such parallels exist in historical developments. A particular change cannot happen so that a long-term goal is reached. Rather, an apparent goal-directedness emerges from the fact that a constant pressure on the replication process, including pressure from co-evolving replicators, results in the survival of the replicators that are successively better adapted to that pressure. The goal-directedness arises in the eye of the observer, who sees the lineage of replicators going back in time from the currently existing forms, to the exclusion of all the alternative paths, which have not survived to the present day, or have not left any records at all.

What motivated Lass (1974) to propose ‘orthogenesis’, is that he saw pure coincidence as the only alternative, which he deems unsatisfactory. However, just as goal-directedness has been eliminated from biological theorizing without sacrificing the possibility of accounting for seemingly teleological developments, so is the case with linguistic historical conspiracies.

If a certain structure is selected against, it will be cumulatively diminished over successive generations, without the goal of eliminating it being an agent of this change. Since, as is argued here, quantity oppositions were not selected for because their expression was increasingly unreliable, then each change along the way can be thought of as being driven by the constant pressure, i.e. the unreliability of quantitative contrasts.

As for the interrelatedness of the diminishing role of quantity and GVS, an explicit link between these was made by Trnka (1982) and taken up by Perkins (1977).³³ Trnka suggests that the trigger for GVS was the reinterpretation of the quantitative opposition in high vowels. Under his account, the short counterparts of the long high vowels, that is /u/ and /i/ were lax so that they came to form length oppositions with /e:/ and /o:/ rather than /i:/ and /u:/, and the long counterparts were “either very close or slightly diphthongal”. Afterward, as /e:/ and /o:/ were phonologically coupled with /u/ and /i/, they were raised to match them also phonetically. A major weakness of this account is that it rests on an early dating of the laxing of /e/ and /o/. The orthoepist evidence adduced by Lass (1999: 88) clearly shows that /i/ and /u/ were lax to /ɪ/ and /ʊ/ toward the end of the seventeenth century, and so this change could not have played a role in the first stages of the GVS. The insight that the diminishing role of vowel length is linked to the advent of shifting, however, is taken over by the account to be presented in the following.

3.4 EXAPTATION OF QUALITATIVE DIFFERENCES

Bringing together the issues presented so far, a first approximation at an account, involving many key arguments, is presented in the following section. This presentation, employing the concept of ‘exaptation’, involves a critical discussion of the usefulness of this concept for accounts of phonological change.

As already mentioned, the main hypothesis put forward by this thesis is that the increased likelihood of English vowels to undergo vowel chain shifts is a consequence of the decrease in the role of duration for the maintenance of phonemic oppositions. If the replacement of

³³ Ostensibly through Stampe (1972), though the latter mentions neither Trnka (1982) nor vowel length.

duration-based oppositions by quality-based oppositions is seen as a co-optation of non-functional qualitative differences inherent in length-based phonological oppositions of Middle English vowels to the function of the expression of vowel contrasts, it can be thought of as ‘exaptation’. Exaptation is a concept borrowed into linguistics from biology, and it has attracted a lot of attention within the linguistic community. The popularity of this concept is testament to the renewed interest in making use of similarities between linguistic change and biological evolution. Its fate is also illustrative of the pitfalls of borrowing terms from other domains without making clear what their applicability in the target domain might actually be. In the following, the biological concept of exaptation, its borrowing into linguistics and its possible application to the history of English vowels is presented. It is argued that exaptation in linguistics is no more than a label, albeit a useful one, whose application is only justified when language is seen as a truly evolutionary system. Its usefulness for the scenario at hand lies in highlighting the growing role of qualitative differences for the maintenance of phonemic vowel distinctions as a reaction to the diminishing reliability of duration to do so. Thus, the nexus of a set of rhythmically motivated quantitative changes on the one hand and qualitative changes on the other hand can be brought into focus.

The term exaptation was coined in an influential paper by Gould and Vrba (1982) to refer to “features of organisms [which are] are non-adapted, but available for useful cooptation in descendants”. Examples that the authors give include mammalian sutures as an exaptation for parturition, birds’ feathers and sexual mimicry in hyenas. The first of these examples was discussed by Darwin (1859: 197) himself and illustrates the concept very well. The existence of joints in mammalian skulls facilitates birth. However, they cannot be labeled as adaptations for facilitating birth, since they are also present in birds and reptiles, which hatch from eggs and so easing birth could not have been a driving force in the genesis of the sutures. The joints did not arise because they facilitate birth in reptiles and yet they now do so in mammals, so a disconnect between genesis on the one hand and current utility on the other is particularly clear here. Another example worth mentioning is the case of sexual mimicry in hyenas, because it illustrates the second of the two possible sources of exaptation, a point which will be discussed later on. In the spotted hyena, female genitalia resemble very closely male genitalia. This is very useful in greeting rituals which are essential in identifying individuals returning to the pack. According to Gould and Vrba (1982: 9), it is plausible to

argue that this mimicry of male genitalia by female genitalia is not an adaptation to the greeting ritual, but that they are “automatic, secondary by-products” of the high levels of androgen found in female hyenas. To return to the issue of how exaptations come into being, Gould and Vrba (1982: 12) give two possible sources of exaptations, namely “adaptations for another function” and “non-aptive structures”. In other words, structures that were adaptations for some function can (additionally) be exaptations for another one (as is the case with mammalian sutures) or, alternatively, structures that had no function at all can be co-opted to have one (as might be the case with sexual mimicry in hyenas).

The formation of this new term was seen by the authors as a step toward undermining the purely adaptive view of evolution, which is just one instantiation of Gould’s ideologically motivated anti-adaptationist agenda (cf. Dennet 1995: Chapter 10). Gould and Vrba (1982) claimed to have identified a mechanism of evolution which provides an alternative to the hegemony of adaptation as its only driving force. Since, they argue, features which were not honed by natural selection can contribute to the fitness of an individual, adaptations can no longer be seen as the only way to explain change in species. In contrast to this estimation of the importance of exaptation, it is argued persuasively by Dennett (1995: 267) that the addition of exaptation to the conceptual inventory of evolutionary biology might be seen as nothing more than an introduction of a useful term into the purely adaptive view of evolution. As he points out (1995: 281), “every adaptation is one sort of exaptation or other” since “every adaptation has developed out of predecessor structures each of which either had some other use or no use at all”. Since no function lasts forever, changes in function are inevitable, and this results in structures being exposed to different selective forces than originally, and corresponds to exaptations having their source in structures evolved for another function. The only other logical possibility for an adaptation to arise is for selective forces to start selecting for a feature that previously had no function at all, and this would correspond to the second type of exaptations, namely to those originating from ‘non-aptive’ structures. Hence, each adaptation must start as an exaptation, and exaptation is nothing more than a useful term to refer to the initial stage in the lifetime of adaptations.

Regardless of the rather modest contribution that exaptation can make in conceptualizing biological evolution, the term has been borrowed into linguistics, where its possible relevance has been considered from various angles over the past twenty years. There have

been attempts to apply it within various research programs, which can be divided roughly into two groups. First, there have been attempts, some of them rather famous, to bring exaptation to bear on discussions of the emergence of (components of) the language faculty. Some scholars still want to see language as such as an exaptation, i.e. as having arisen irrespective of any adaptive value (Hauser, Chomsky & Fitch 2002; Boeckx & Piatelli-Palmarini 2005), a view shown to be ill-founded by Pinker and Bloom (1990). These accounts seem to place much stock in the non-adaptive, revolutionary aspect of exaptation as originally postulated by Gould and Vrba. MacNeilage (1998) proposes an account under which phonology is an exaptation of the cyclical properties of the mandible. Samuels (2011), also with regard to phonology, argues that it is an exaptation from other cognitive domains. Carstairs-McCarthy (1999) suggests that phrase structure was co-opted from syllable structure. In contrast to these applications which seek to throw light on the emergence of (components of) the language faculty, there have also been attempts to apply exaptation to the study of change within particular linguistic systems, that is within the domain of historical linguistics. These are of much more interest here, as it is the culturally evolving system of a language which is investigated in the present account. Actually, this is the area where linguists', more specifically Roger Lass (1990) first used exaptation. In a paper, titled "How to do things with junk", he postulates the interpretation of certain morphosyntactic developments in Germanic as exaptations. Curiously, he decided to narrow down the scope of exaptation by limiting it only to the exaptations of the second type as originally postulated by Gould and Vrba, that is to structures that are absolutely non-adaptive, the eponymous "junk". Whereas Gould and Vrba (1982) see non-adaptive features as automatic by-products of other features of the organism, Lass (1990) defines them as characters which were once adaptive, but have lost their function. Vincent (1995) argues that linguistic exaptation does not operate on junk, and Lass (1997) agrees. Still, the issue of junk has remained part and parcel of discussing exaptation within historical linguistics; for example Giacalone Ramat (1998) brings it up once again, if only to criticize it. The debate of the concept in linguistics has since focused mostly on assessing the relationship between exaptation and other mechanisms of morphosyntactic change, such as reanalysis, analogy and grammaticalization. And so Taught (2004) concludes her discussion of this relationship by claiming that a sufficiently narrowly circumscribed exaptation might indeed refer to a hitherto unnoticed

type of morphosyntactic change, which, even if infrequent, is still real. De Cuypere (2008) is much more pessimistic about the usefulness of the concept, and, having criticized Lass's characterization of exaptation as too narrow, admits that the broader conceptualization, more in line with the biological interpretation of the term, is too all-encompassing to be actually helpful. All those discussions continue the metaphorical understanding of the link between biology and linguistic systems represented by Lass (1990) (but not Lass 1997). They reject out of hand the possibility of a technical application of evolutionary thinking to language, as is usually the case (this relates to Blevins 2006, as well as her commentators, be it critical or approving, e.g. Hamann 2006, Kiparsky 2006 and Smith 2007). Consequently, they boil down to discussions about the usefulness of a biologically inspired metaphorical label and miss out on the benefits of seeing linguistic evolution as a strictly evolutionary process, as will be illustrated in the following discussion of the developments in the history of English vowels. Admittedly, the concept has been discussed in linguistics from an evolutionary perspective, in a paper by Croft (2002). However, since he assumes a rather peculiar view of generalized Darwinism, where random mutation and adaptation are not seen as its essential components, it is in the end unclear what role he assigns to exaptation. Since he claims on the one hand that "there is no logical inference from the generalized theory of selection that requires language to display [...] exaptation" and on the other hand that "[exaptation] ha[s] [a] clear analo[g] in language, albeit with different causal explanations particular to the domain of language" (Croft 2002: 86), it is difficult to say what criteria he proposes be applied to decide whether one is dealing with exaptation or not. At any rate, since adaptation through natural selection is in fact an indispensable mechanism of neo-Darwinian evolution, Croft's arguments do not have any far-reaching implications for the present discussion. The development to be described is one where the exaptation, that is a change in the selection pressures to which a given structure is sensitive, is a clearly identifiable stage in the development of the English vowel system. The relevant events are presented below.

In Old English (Figure 8), all the way through Middle English (Figure 9), there were phonemic vowel contrasts that were maintained by phonological length, whose acoustic correlate was phonetic duration. A comparison with a Present Day English variety, RP (Figure 10), reveals that there is no vowel pair where length alone would be decisive in maintaining the contrast. The differences between the RP system and the ME system are a

result of a whole host of vocalic changes that have taken place from the EModE period onwards, with the first important changes affecting the high-mid vowels, long and short, as presented in Table 3.

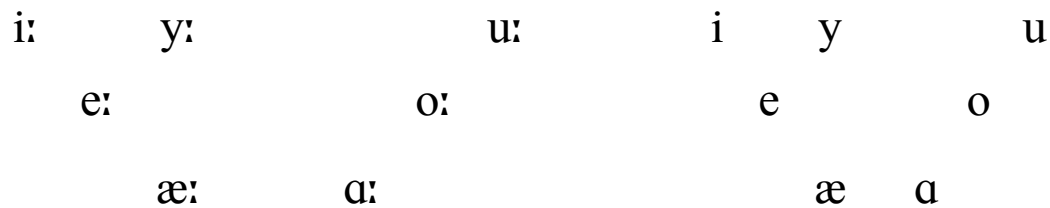


Figure 8: OE pure vowels, 10th century, after Hogg (1992a: 85)

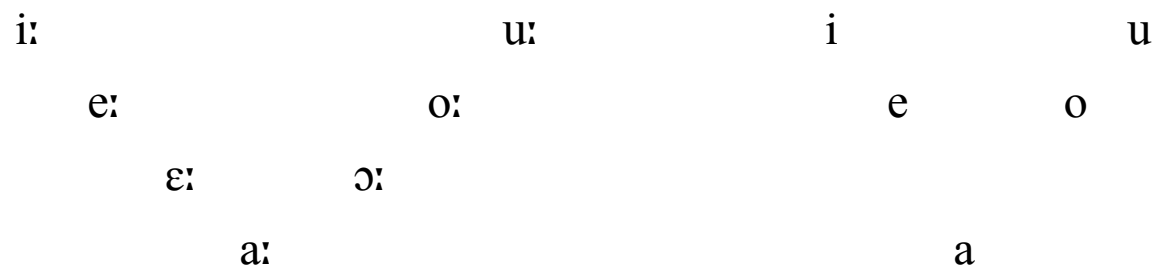


Figure 9: ME pure vowels, 15th century, after Lass (1999: 68)



Figure 10: RP pure vowels, after Lass (1999: 69)

	15th century	16th century
meet	e:	i:
met	e	ɛ
boot	o:	u:
pot	o	ɔ

Table 3: Changes to the high-mid vowels in EModE (the long vowels after Lass 1999: 72, the short vowels after Lass 1999: 91)

What can be said about the fate of these four vowels occupying the high-mid region in the 15th century is that the long vowels raised (as part of the GVS) and the short vowels lowered (as part of the Short Vowel Shift). Additionally, in place of a two-way qualitative difference further differentiated by a length contrasts, these four lexical sets now possess four distinct vowel qualities, albeit still distinguishing between long and short.

Now, what suggests the exaptive nature of these changes is the observation that in languages which employ length as a feature maintaining vowel contrasts, purely quantitative differences tend to be accompanied by qualitative differences, particularly for the high and mid vowels³⁴. This is for instance the case in Czech (Dankovičová 1999), Hungarian (Szende 1999), Slovak (Hanulíková & Hamann 2010), Swedish (Engstrand 1999), and Zurich German (Schmid 2004).³⁵ An analogous situation has been reconstructed for Classical Latin (Loporcaro 2011: 110), Sanskrit (Bloch 1965: 35) and early Proto-Slavic (Schenker 1993: 79). Aside from the synchronically observed correlations between length and peripheralness, diachronically speaking, when quantity-based contrasts are re-interpreted as quality-based contrasts, then short vowels acquire quality that is less peripheral, or more ‘reduced’, e.g. in the developments Old Icelandic > Modern Icelandic (Þráinsson 1994), Sanskrit > Hindi (Masica 1993: 35), early Proto-Slavic > late Proto-Slavic (Sussex & Cubberley 2006: 33), Latin > Common Romance (Loporcaro 2011: 115). Admittedly, the exact source of these qualitative differences between phonologically long and short vowels is not entirely clear (see discussion in Section 4.5.2.1 below). In fact, short vowels could be expected to be higher, since, other things being equal, low vowels are longer than high vowels (House & Fairbanks 1953; Peterson & Lehiste 1960; House 1961). The ‘the lower, the longer’ generalization, however, holds only within the classes of long and short vowels. As the above-cited sources indicate, when pairs of qualitatively similar vowels which differ in length are compared, the obverse is true, that is the lower member of the pair turns out to be shorter. In short, within the long and short sequences, the lower equals the longer, but across classes, shorter equals lower. It does not seem too far off to suspect that the same was true of

³⁴ The qualification ‘particularly for the high and mid vowels’ is in agreement with Donegan’s (1978: 64) argument that low vowels are not subject to tensing, that is that the qualitative differences in this region are due to height or length, but not to tenseness.

³⁵ The vowel pairs of Dutch (Gussenhoven 1999), described as distinguished by tenseness, are characterized by the long counterparts having a more peripheral, and diphthongal, realizations, not unlike in PDE.

Old English vowels. If that indeed was the case, then the primarily durationally expressed contrasts between phonologically long and short vowels can be speculated to have been accompanied by qualitative differences, with the short members of the pairs being lower relative to the long members. Consequently, there was a character – the accompanying qualitative differences – which was not functional. Since in OE and ME the contrasts were phonetically realized first and foremost by duration, the qualitative differences can be said to follow from them and so constitute a kind of a by-product. This by-product was later harnessed to do the work of maintaining phonological contrasts, as is the case with the developments of the high-mid vowels specifically, but also in the case of the overall development of the vowel system from EModE to PDE generally. A development in which a by-product becomes functional seems to bear the hallmarks of exaptation. The qualitative differences are not leftovers from a previously functional system, nor does their rise to the status of correlates of phonemic oppositions represent a complete novelty. After all, quality always plays a role in distinguishing vowel phonemes, it is just that a kind of contrasts has arisen, when qualitative differences team up with distributional characteristics to create the ‘tense’ versus ‘lax’ vowels of Present Day English. Those two facts, i.e. that the qualitative differences are no ‘junk’ and that their cooptation does not result in a complete innovation would disqualify them as exaptations in Lass’s original formulation. But since neither junk status nor novelty are necessary properties of exaptations (Dennett 1995; Lass 1997; Traugott 2004), co-opted qualitative differences do remain legitimate candidates for the status of exaptation.

Assuming that the replacement of length with tenseness as the primary feature distinguishing two classes of English vowels is an exaptation of inherent qualitative differences, two analyses are presented below: one superficially biological, simply borrowing ‘exaptation’ as a term, and the other a strictly evolutionary one, borrowing the concept of exaptation within an adaptive evolutionary framework. An analysis which takes exaptation to be a biologically inspired metaphor could look like the following. The long members of the pairs, since they are pronounced with longer durations possess slightly more peripheral qualities, which means, in the case of high-mid vowel higher and fronter and in the case of the high-mid back vowel higher. Their short counterparts, which are by definition constantly pronounced with shorter durations, possess less peripheral, that is lower, and, in the case of

the front vowel, more central qualities (Stage 1 in Table 4). The qualitative differences become at some point co-opted for the maintenance of the distinctions. Once exapted, the qualitative differences are further exaggerated (Stage 2 in Table 4).

	Stage 1	Stage 2
meet	[e:]	[i:]
met	[ɛ]	[ɛ]

Table 4: High-mid vowels before and after exaptation of qualitative differences, version 1

Such an account is reminiscent of what since 1976 has been known as phonologization (Hyman 1976). It is a development where an ‘intrinsic’ phonetic effect becomes ‘extrinsic’, that is, in Hyman’s (2008: 385) terminology, part of language specific phonetics, and then, finally phonological. In Hyman’s (1976: 33) words it is a case where “an intrinsic byproduct of something else, predicted by universal phonetic principles, ends up as unpredictable and hence, extrinsic”. Thus, identifying the development as an exaptation in the metaphorical sense does not go beyond a traditional phonological account. In fact, it could be seen as a step back in a sense, since to be able to call a development an instance of phonologization one would have to be very precise about the way in which the qualitative differences are intrinsic phonetic effects. For exaptation, however, the origin of the structure is essentially irrelevant.

An account invoking exaptation within a strictly evolutionary conception of language, on the other hand, does go beyond the phonologization account. First of all, it recognizes that co-opted features must have an advantage over resident features to replicate successfully. Here, it was posited that the new contrasts were better able to express phonological contrasts, and this is what helped them oust the residents in the long run. This adaptive aspect of exaptations is likely to be overlooked on the metaphorical reading, as in that case the label is taken over from the source domain (here, with the now debunked idiosyncrasies ascribed to it by Gould and Vrba 1982) and is not updated. At any rate, acknowledging the adaptive advantage of the new structure forces the question of why it replicates better than the original structure. An answer to this question will in effect provide an answer to another question, i.e. why did the transition from Stage 1 (‘intrinsic’ qualitative differences) to Stage 2 (‘extrinsic’

qualitative differences) happen. In other words, it will provide an answer to the actuation problem. Additionally, the further exaggeration of the qualitative differences is only to be expected under this account. A strictly evolutionary account, since it holds the promise of answering the first two questions and of providing a coherent attempt to provide a strictly evolutionary account of the development is outlined in the following.

There are reasons to believe that in ME short vowels could be lengthened and long vowels could be shortened for rhythmic reasons (Ritt 2012: 404). The preference for the trochaic foot, not only within words, but also on the utterance level, spanning several lexical items, must have been operative. Recalling that feet are good candidates for replicators, it can be postulated that the trochaic foot was thriving in ME. Since replicators constitute an important part of the environment for other replicators, and since rhythm is acquired early on in first language acquisition, the trochaic foot can be thought to have exerted a considerable pressure on other replicators, including vowel phonemes. The coexistence of phonemically long and short vowels, together with the preference for the trochaic foot could have led, phenotypically, to the formation of ambiguous forms, as presented at Stage 1,5 in Table 5.

	Stage 1	Stage 1.5	Stage 2
meet	[e:]	[e']	[i:]
met	[e]	[e']	[e]

Table 5: High-mid vowels before and after exaptation of qualitative differences, version 2

With such pronunciations, duration was no longer a reliable cue to the identity of the respective vowels. At the same time, the phonemes were already slightly different in terms of their qualities, and so these were likely to be paid attention to by language learners, who would have arrived at grammars producing quality-based contrasts, with only a secondary difference in length. A further exaggeration of the qualitative difference would have resulted in the values evident at Stage 2. In effect, the development was advantageous for both the vowel phonemes (after the exaptation of the qualitative differences they maintained their identity which was threatened before) and for the trochaic foot, which could now be expressed more robustly, once duration was less tied up in the expression of phonemic contrasts.

Since evolutionary change acts without foresight, it is no wonder that the newly established vowel qualities found themselves in the same sort of durational relations with already existing vowels as they had been before the change. This time, with the durational differences weakened, the situation was even more precarious. What can be observed in the transition from the 16th to the 17th century is the replacement of four vowel contrasts expressed mainly by duration with contrasts that are qualitative in nature (see Table 6).

	16th century	17th century
meet	i:	i:
bit	i	ɪ
beat	ɛ:	e:
bet	ɛ	ɛ
top	ɔ	ɒ
home	ɔ:	o:
moon	u:	u:
put	u	ʊ

Table 6: 16th century quantitative oppositions replaced with qualitative oppositions in 17th century (the long vowels after Lass 1999: 72, the short vowels after Lass 1999: 91)

To recapitulate, exaptation can be a useful concept in analyzing phonological change, but only insofar as it is used within a strictly evolutionary view of language. The metaphorical understanding only provides a convenient label, which, in this particular case, does not add anything to an account that would not employ it. Applied within an evolutionary approach, on the other hand, it forces the ‘why’ question, an answer to which must lie in the increased replicative strength of the new structure over the old structure with respect to the environment, including other replicators. An attempt to provide such an answer was presented above. Its plausibility can be better assessed when formalized clearly, a full account fleshing out these ideas and formalizing them in OT is presented in Chapter 4 below.

4 THE PROPOSED ACCOUNT

4.1 INTRODUCTION

This chapter proposes an account of how vowel quality has become the locus of change in English in response to a weakened role of duration as a correlate of the phonemic length oppositions, which contributed to the increased rate of vowel shifts. It is suggested here that vowel duration has become unreliable as a marker of phonemic oppositions as a result of a number of historical developments. The first of them is the fixing of lexical stress, an event that took place already in Germanic. The second is the continuing shift in the direction of stress-timing (or word-rhythm), which has been taking place since Old English times, through the various quantitative adjustments dubbed together as ‘the length conspiracy’ by Lass (1974), and by consonantal simplifications, too. The third development that contributed to the weakening of the reliability of duration as a cue for phonemic vowel oppositions is, indirectly, the reduction and loss of unstressed vowels. All those events taken together, having diminished the reliability of duration as a marker of vocalic phonemic oppositions, resulted in the exaptation of the qualitative differences between vowel pairs previously distinguished by duration, which set in motion the chain shifts observable in a number of varieties of English today. As a result of these developments, the number of distinctive vowel qualities rose, qualitative variation has ceased to be easily predictable from quantitative variation, and hence become salient, so that very small qualitative differences are paid attention to by listeners. Stress placement is still relatively stable (though not as in OE, since there are derivational affixes that shift stress, i.e. that eliminate the primary stress of one form altogether), lengthenings and shortenings are not that available to be exploited for social purposes, since duration is influenced by a number of factors (prosody, coda sonority). Since the social need toward exploiting variation, which is the motivation behind change, is still there, change has as its locus qualitative differences.

The following is the presentation of this development in evolutionary terms using the formal apparatus developed in OT. As mentioned before, the quest for the causes of the Great Vowel Shift is recast here as a quest whose goals are both more general and more modest than those of many previous attempts. In the past, the goal was usually to pin down the cause

of this particular shift. Here, it is argued that English has become a vowel-shifting language, and that the GVS, along with later vocalic shifts, are a part of this broader development. In other words, it is believed that English has in general become more likely to undergo vowel shifts than it used to be before Early Modern English times. The goal of this thesis is therefore rather different than finding a cause (or ‘the’ cause) which set off the GVS, that is one particular ‘event’, since it tries to provide an account of a global change in the language, i.e. of the acquisition of a new macro-property of instability of vocalic qualities rather than a one-shot reorganization of the long vowel system. It is, however, at the same time also more modest, in that no deterministic account is provided which could make predictions about developments that must necessarily take place in other languages or in the same language at other stages in its development given a certain set of initial conditions and general laws. Instead, it merely shows how individual adaptations took place, and argues that they happened according to evolutionary principles. The goal is not to show that the GVS must have happened, or that English must have become a vowel-shifting language, anymore than it is the goal of an evolutionary account to show that biological evolution had to result in the development of any particular species (cf. the vacuity of the misconception that properly constructed evolutionary biology should entail the necessity of the development of modern homo sapiens). Both when providing an account of the evolutionary history of an extant species and when drawing up an evolutionary account of the development of an extant language, the starting point of an investigation is the entity at hand, but the goal of the account is not to show that it must have come into existence, but rather, given that it has, what were the individual stages and how they were motivated.

4.2 OLD ENGLISH

The entities which underwent the GVS in the EModE period and which have been undergoing qualitative shifts at an increased pace ever since are the vowel phonemes of English, and the vowel phoneme systems that they have formed throughout history. The histories of vowel phonemes, and of vowel phoneme systems, starting from OE times through to PDE are told in the following sections.

Two major differences can be observed between the vowel systems of OE and of PDE. One is that the qualitatively comparable vowels of the two systems are distributed differently

in the lexicon. This is the result of vowel shifts, splits and mergers, and could be seen as a group of events that affected individual vowels, or, at most, subsets of the vowel system. The other major difference is that the vowel system where phonemic contrasts between two classes of vowels were expressed by duration (phonologically: length), has now become superseded by one in which these contrasts are expressed primarily by quality (phonologically: tenseness). This difference has to be seen as stemming from a reorganization which has affected the entire vowel system of the language. In the following, it is suggested that the latter has contributed to bringing about the former, and that both have been influenced by developments in the suprasegmental domain.

With regard to the segmental level, it is widely acknowledged that (as already illustrated above in Section 3.4) the OE vowel system was marked by the presence of length oppositions. The Classical OE (that is, tenth century West Saxon) vowel system most likely comprised seven vowel pairs which were distinguished phonologically by length and phonetically by duration, and, as Hogg (1992a: 85) asserts, “there is no consistent evidence of any corresponding qualitative shift” (where by ‘qualitative shift’ he refers to a qualitative difference between the long and short members of the seven oppositions, of the sort observable in PDE *bead* / *bid* pair, for instance). This very symmetrical system was soon to be reorganized, and the reasons for that are to be sought in the developments in suprasegmentals. The seeds of those developments had been sown even before OE started.

The first of the developments suggested to have ultimately contributed to the weakening of the reliability of duration as a cue toward maintaining phonemic length contrasts is the fixing of lexical stress in Germanic. Through enabling the accumulation of segmental effects of stress, its fixing contributed to the weakening and loss of unstressed vowels, which ultimately led to an increase in the number of stress clashes and lapses. Their avoidance then became one of the reasons why duration became more variable, i.e. less tightly bound with the expression of phonemic oppositions.

4.2.1 *Stress assignment in OE*

OE stress assignment is a continuation of stress assignment in Proto-Germanic, which was innovative with respect to Indo-European. In Germanic, “free” word stress of Indo-European was “fixed” so that it “invariably occurred on the root syllable, which in the vast majority of cases was the first syllable of a word” (Lahiri, Riad & Jacobs 1999: 336). There is disagreement as to whether morphology was at play, that is whether reference to root has to be made to derive generalizations about stress assignment in Proto-Germanic, or whether phonology alone was able to assign stress, that is whether stress was word initial, not root initial³⁶ (for a presentation of the two “schools of thought”, see: Lahiri, Riad & Jacobs 1999). Regardless of the situation in Proto-Germanic, though, in Old English, it was not phonology alone, but an interplay of morphology and prosody which determined the location of stress, and the fixed stress in OE fell on the leftmost syllable of the root (Hogg 1992a: 99; Lass 1992a: 85; Minkova 2006). As a result, OE is described as having ‘fixed lexical stress’, since for each phonological word the stress was ‘fixed’ on one and the same syllable and it could not ‘move’ to signal meaning differences, as it used to be the case in Indo-European, and evidenced for example by the Greek examples: *patér* ‘father-NSg’ vs. *páter* ‘father-VSg’ (2007: 79). To refer to two major kinds of stress-assignment systems laid out by Apoussidou (2006: Chapter 2), a stress system like the one of Germanic would be described as grammatical stress, rather than lexical stress, since in a system such as that in OE stress is assigned by the grammar (where grammar, naturally, includes phonology), and no stress marking is necessary on the lexical level. From now on the terms ‘grammatical stress’ and ‘fixed lexical stress’ are used interchangeably.

It is also argued, that apart from the placement of stress, also its character changed on the way from PIE to Germanic, namely that pitch was replaced by duration and intensity as the main correlates of stress (Iverson & Salmons 2003). The role of duration as a correlate of stress is a prerequisite for the increased durational differences between stressed and unstressed vowels, which played a part in the reduction of unstressed vowels and in strengthenings of stressed vowels.

³⁶ The latter position seems to be adopted, without comment, by Harbert (2007: 79), who states that “[t]he GMC languages, like other Western European languages, have prosodically fixed word stress – that is, stress whose position is (mostly) predictable on the basis of the phonological shape of the word, and does not depend on grammatical information”.

Since in OE stress is assigned by grammar, stress assignment can be modeled by making use of the relevant metrical constraints. Two basic constraints with regard to metrical structure of languages relevant here are $RH\text{TYPE}=\text{TROCHAIC}$ and $RH\text{TYPE}=\text{IAMBIC}$, introduced by Prince and Smolensky (1993 [2002]: 56) on the basis of the generalization that for each language, either trochaic or iambic feet predominate, rather than some mixture of the two types of feet. Their relative ranking in a given language (alternatively, under the constructionist view on constraint emergence, the construction of one or the other constraint in the learner's grammar) decides whether trochaic or iambic feet will characterize surface forms of words³⁷.

1) $RH\text{TYPE}=\text{TROCHAIC}$ (TROCHAIC): The leftmost syllable in a foot is the head syllable

The third constraint utilized in the tableau is FTBIN , also introduced in Prince and Smolensky (1993 [2002]: 50), as an OT constraint version of the requirement of binarity of feet proposed in McCarthy and Prince (1986)³⁸, and which requires that feet are binary at some level of analysis, either at the syllable, or mora level.

2) Foot Binarity (FTBIN): Feet are binary at some level of analysis (μ, σ)

Following Minkova (2006), syllables are assumed as the basic component parts of feet in OE. FTBIN makes sure that degenerate, one-syllable feet are excluded from the analysis. Thus, a simplified scenario for stress assignment in OE lexical disyllables is presented in Tableau 2 below.

³⁷ In Minkova's (2006) account of stress assignment in OE these are made superfluous, since her model yields trochees simply by requiring that left edges of roots coincide with heads of feet, thus trochees are emergent from the interplay of prosody and morphology, doing so by recourse to alignment constraints of McCarthy and Prince's (1995b) 'Prosodic morphology'. Though it is particularly well-suited to handle prefixation, the benefit of this approach pales in comparison to the generality of $RH=\text{TYPE}$ constraints. These, in combination with an alignment constraint requiring that left edges of feet coincide with left edges of roots can ensure that prefixes are not stressed.

³⁸ which Prince and Smolensky (1993 [2002]: 50) suggest ultimately derives from Prince (1980)

Underlying: σσ e.g. <i>wuldor</i> ‘glory’	FTBIN	TROCHAIC
𐌿𐌺 (ós)		
(só)		*!
(ó)(ó)	*!	

Tableau 2: Stress assignment in two-syllable OE words

OE lexical words were mostly disyllabic. In Getty’s (2002, cited in Minkova 2006: 102) study, only 23,9 percent of words were lexical monosyllables, and 4,6 percent were words of three syllables, which must mean that a vast majority were disyllables.

Rare though they were, lexical monosyllables did occur though, and their stress patterns have to be accounted for. Minkova (2006) assumes that they contained an empty element in place of the second syllable required by FTBIN. If FTBIN is to refer to syllables rather than moras in OE, the presence of some such empty element, referred to as ‘catalexis’ seems unavoidable. Tableau 3 presents how it can be modeled in OT.

Underlying: σ e.g. <i>scip</i> ‘ship’	TROCHAIC	FT-BIN
𐌿𐌺 (ó. ø)		
(ó)		*!
(ø. ó)	*!	

Tableau 3: Stress placement in OE lexical monosyllables

Although three syllable and four-syllable words did exist in OE, a vast majority of them could arise only in derived environments. Inflexional suffixes never bore stress and can be seen, just as prefixes, as invisible to stress assignment. Derivational suffixes did not bear main stress either, but they could be endowed with secondary stress. In words with two stresses, then, it has to be resolved which of the two stresses is primary. Since the primary stress was always the leftmost one, no marking of affixes with regard to their influence on stress placement has to be assumed for OE. Rather, the placement of the main stress was resolved by the grammar. To make sure that in longer words it is still the first syllable of the root that bears the main stress, two other constraints must be invoked, namely MAIN-LEFT and MAIN-RIGHT, where the former requires that the leftmost foot in a word be the head foot, and the latter that the rightmost foot in a word be the head foot (Tesar & Smolensky 2000,

cited in Apoussidou 2006: 19). Consequently, words consisting of four syllables will be stressed as presented in Tableau 4.

3) Constraints on the placement of main stress

MAIN-LEFT: The leftmost foot in a word is the head foot

MAIN-RIGHT: The rightmost foot in a word is the head foot

Underlying: σσσσ e.g. <i>gafolræden</i> ‘tribute, rent’	TROCHAIC	FTBIN	MAIN-LEFT	MAIN-RIGHT
$\mathbb{E}/(\acute{\sigma}\sigma)(\grave{\sigma}\sigma)/$				*
$/(\grave{\sigma}\sigma)(\acute{\sigma}\sigma)/$			*!	
$/(\acute{\sigma})(\acute{\sigma}\sigma)(\acute{\sigma})/$		*!*		

Tableau 4: Stress assignment in four-syllable OE words

The details of the influence of morphology on stress placement are not elaborated on here. A detailed discussion can be found in Minkova (2006). Some very important characteristics have to be mentioned, though. Suffice it to say that root-initial stress in OE persisted even in almost all derived forms. Granted, there were a few stressable nominal prefixes such as *and-*, the stressable counterpart of *on-*, as in *andgiet* ‘understanding’. However, since in such words the root syllable was also stressed, such words can be seen as compounds rather than prefixed forms, and true prefixes, such as *on-* were never stressed (Lass 1992a: 85). Furthermore, derivational suffixes never shifted the main stress away from the first syllable of the root demoting it to secondary stress (as is the case for example in PDE ‘*absent* – ,*absen*’*tee*), nor did they move the main stress altogether (as is the case in PDE ‘*symbol* – ,*sym*’*bolic* or *re*’*pute* – ‘*reputable*) (Minkova 2006: 98). From OE through ME, then, lexical stress was much more fixed than it is nowadays.

In view of the further developments, it is important to note here that OE words were on average longer than in later periods, with an important aspect being the lower number of monosyllables than later. After the erosion of morphological endings and final schwa loss, many words became shorter by a syllable than before. Thus, relatively speaking, each OE stressed syllable was much more likely to find itself in a context where the trochaic foot could be realized than it has been the case in later periods. For example, in disyllabic words,

their stressed vowels were parts of binary feet regardless of other words around it. The same is true of trisyllabic words, where the third syllable was a suffix, and therefore extrametrical. Due to a higher proportion of disyllables to monosyllables than later on, the rate at which each vowel appeared in a monosyllable was lower in OE than later on. In other words, the stressed syllables of head feet received the same amount of stress, and were more independent of the words that preceded or followed them. After schwa loss, and with the substantial increase in the number of monosyllables, this situation would change and the likelihood that a vowel is part of a binary foot would come to depend not only on the properties of the words it is a part of, but also on the properties of the words which precede and follow those words.

This development, that is the fixing of stress on the first syllable of the root, is important for vocalic developments in two disparate ways. On the one hand, fixed lexical stress contributes to the entrenchment of prominence in the stressed segments. Thus, the presence of fixed stress enhances the stress-timing appearance of a language (more is said on the drift toward an ever-more stress-timing language type in the next section). On the other hand, fixed lexical stress can contribute to repeatedly occurring stress clashes and lapses, especially once unstressed vowels are reduced and then lost, as a result of the first effect, that is of the accumulation of segmental effects of stress. Thus, the presence of fixed stress undermines the stress-timing nature of a language by standing in the way of rhythmic alternation. The mechanism behind the former, and the consequences of the latter are as follows.

4.2.2 Fixed lexical stress as a factor enhancing the stress-timing of English through enabling the accumulation of segmental effects of stress

To begin with the first of those effects, if one and the same vowel is always stressed in a given word, the phonetic manifestation of stress can cumulatively increase. This leads to an increase in prominence between stressed and unstressed syllables, and such a contrast is a typical feature of languages traditionally classified as stress-timed. On a segmental level, it can be seen in the disappearance of vowel length distinctions in unstressed position which took place on the way from Germanic to Old English (Lass 1992a: 77).

Since expressing stress is a matter of relative, not absolute, prominence, stressed vowels have greater loudness, duration, and higher (or lower) pitch (or a combination of these three properties) than unstressed vowels (Nootheboom 1997). If it is always the same vowel that is stressed within a word, then that vowel is always more prominent relative to the unstressed vowels in that word. In line with the views of Paul (1880 [1920]) (discussed above in Section 2.4.2) and in their modern version, of, among other researchers, Bybee (2001) and Pierrehumbert (Pierrehumbert 2001, 2002b; Pierrehumbert 2006) (discussed in Section 2.4.6 above), the effects of language use can become entrenched to influence language structure. *Pace* Bybee (2001), the accumulation of the repeated effects of language use does not have to mean that each word is stored as an undivided sound image, and that categories such as phonemes do not exist, though. Such a view is ill-equipped to account for developments such as vowel chain shifts, where the phonemic category of a vowel as a whole changes, as opposed to a select group of words. Reconciling the existence of word-specific, frequency-driven phonetics with the reality of categories, Pierrehumbert (2002b: 107) suggests that during production/perception of speech the abstract categories, such as phonemes, are activated, but the subphonemic details of their realization derive from word frequency value stored together with lexical representations of individual words. This metric of word frequency has a bearing on where on the hyper to hypoarticulation scale the word is situated, and thus influences the exact phonetic quality of its segments. This model accounts for the differing developmental paths of very frequent or very infrequent words, a challenge to the Neogrammarian view of sound change at least since Schuchardt (1885).

The arena in which the importance of word frequency is most often invoked is phonetic reduction, the type of change most extensively investigated by Bybee (2001). Reductive sound changes are often most advanced in the most frequent words that undergo them, and least advanced in the least frequent words. An example of this is the fate of sequences of unstressed schwas and liquids in American English. The word with the highest token frequency, *every* has lost the vowel altogether and has become disyllabic; words of average frequency, such as *memory* or *salary* now contain a syllabic /r/, and words with the lowest frequency such as *mammary* or *artillery* are the most conservative in that they have retained a sequence of /ə/ + /r/ (Bybee 2001: 40), referring to (Hooper 1976). Hence, the influence of word frequency on the progress of reductive sound change gains substantial support.

What has to be added, though, is that this reduction does not affect all segments of the word in an undifferentiated manner. In case of the word *every*, for example, it is the second syllable that has been reduced. It is not surprising that it is not the stressed vowel that has been lost. To account for this unsurprising fact, both Bybee's account and Pierrhumbert's model have to allow for prosodic information influencing production. This additional information about the prosodic structure of a word must be activated before the effects of word frequency kick in.

Specific word frequency of individual lexical items aside, it should still be puzzling why the relevant vowels from Hooper's study are schwas to begin with. It is probably the case that when the relevant lexical items were borrowed into English they were made to match the template of distribution of full and reduced vowels present in English at that time. The question then becomes, how did this template arise in the language, or, more specifically, why did unstressed vowels (with the exception of /i/) come to be realized as schwas in English by the eleventh century (Lass 1992a: 77). A more general question is why does a reduction of this sort arise at all, that is a global trend to replace all vowels in unstressed position with schwas. At any rate, the fact that unstressed vowels came to be realized as schwas suggests that just as frequent words (in terms of token frequency) reduce faster than words of average frequency, frequently unstressed segments (in terms of the consistency with which a given segment is unstressed in the various inflectional forms of a word) reduce faster than segments that are reduced less often. Especially in light of the classic Neogrammarian (lexically abrupt) changes that have been attested, the frequency of the lexical items that it applies to can play no role (which is the case, as Bermudez-Otero (2013 (forthcoming)) points out, in the Northern Cities Chain Shift, or /u:/ and /ou/ fronting). The prosodic structure of the words does influence the change, though, in that it is always vowels in stressed syllables that undergo the change³⁹.

³⁹ In English, however, it seems to be trivially true, because there are arguably no full vowels in syllables which are devoid of stress. The only exception to that is /i/, which has remained contrastive in unstressed position. Recently, however, it has been falling together in schwa in non-final position. In final position, it has raised to /i/ in many varieties, a development dubbed 'the happy tensing' by Wells (1982: 257), who notes that it is difficult to date the origins of this development, but cites sources pointing to its presence in mid-twentieth century. In this case, unstressed position has not resulted in weakening, but the vowel apparently fell prey to the restriction on the appearance of lax vowels in open syllables.

Further, as already exemplified by the two examples of change given above, there are reasons to believe that segmental effects of stress are not limited to reduction only. There clearly are instances of non-reductive sound change (e.g. numerous vowel chain shifts, including the GVS), where the change is limited to prosodically strong positions. Namely, just as segments which find themselves frequently and consistently in unstressed position reduce faster than other segments, segments which find themselves frequently and consistently in stressed positions enhance their prominent characteristics more than other segments.

The theory of strengthenings was developed within Natural Phonology by Donegan (1978), who defines them, according to their function as those ‘processes’ which “increase the phonetic properties of individual segments, making them more perceivable” (Donegan 1978: 21), or “those which optimize or maximize phonetic features of individual segments” (Donegan 1978: 62), and which tend to “increase in a segment a property which it already possesses to a high degree” (Donegan 1978: 61). In Natural Phonology, processes are substitutions reflected in historical change, in synchronic phonologies and in child language. For the purposes of the present discussion, only the diachronic strengthenings are considered. One example of a strengthening process is that of tensing, which, for a given height, increases a segment’s ‘color’, that is its backness/frontness and rounding.⁴⁰ And so a tense PDE /i/ is fronter than ME long /i:/, and a tense PDE /u/ is more back than ME long /u:/, with both enhancing the property that they already possessed, that is their specification for front/back. Another strengthening process is that of lowering, which increases a segment’s sonority. This strengthening process does not maximize a segment’s property, but optimizes it by increasing its sonority through decreasing its height. This was the case in the so-called Open-Syllable Lengthening in English, where, for example, the OE /o/ in *nosu* ‘nose’ was lengthened and lowered to /ɔ:/, and OE /e/ in *beran* ‘bear’ was lengthened and lowered to /ɛ:/ in the thirteenth century (Lass 1992a: 47).

There are a number of reasons which support the hypothesis that strong prosodic position favors strengthenings. For one, it is supported by the existence of synchronic phonological

⁴⁰ By this definition, tensing cannot apply to central unrounded vowels, since they have no ‘color’ to increase.

processes which lengthen stressed vowels. This is for instance the case in Hixkaryana, a language with no phonemic vowel length, but in which all stressed vowels lengthen, as in /ne¹mo:ko'to:no/ 'it fell',⁴¹ but also in English in the different distribution of full and reduced vowels which correlates with stress placement in alternations such as /'ætəm/ ~ /ə'təmɪk/ (both examples from Kager 1999: 146). Such alternations, regardless of one's views on the nature of synchronic phonological processes at hand, or lack thereof, must be the result of phonologization, which is a historical process, and which presumably reflects a cognitive response to mechanical, phonetic facts (under the 'amphichronic' view of phonology of Baudouin de Courtenay 1895; Kiparsky 2006; Bermúdez-Otero 2013 (forthcoming), introduced in 2.4.5 above).

Furthermore, being consistently stressed and long can result in the concentration of the features typical of strong prosodic positions even further. A number of historical developments which point to a linkage between being frequently stressed and long is the tensing of long vowels (in English, German dialects, Scandinavian languages, Classical Latin, and Hindi, cited by Donegan 1978: 65). Further, there is experimental evidence that vowels are strengthened in prosodically prominent positions, which might contribute to vowel shifting (Jacewicz, Fox & Salmons 2006).⁴² Since consistently stressed vowels tend to lengthen and since the lengthened stressed vowels tend to raise and diphthongize, then presumably it is a result of accumulated effects of phonetically driven variation.

To account for the segmental effects of prosody which led to strengthenings, the Pierrehumbert (2002b) model is not sufficient. Whereas reductive changes are driven by word frequency, non-reductive changes are not. It is true that different words are affected by chain shifts to differing degrees. This, however, has nothing to do with their lexical frequency, but rather with the phonological makeup of the lexical items at hand, as evidenced by Labov's (1994 [2010]) investigation of the Northern Cities Chain Shift, for instance. One possibility of accounting for the accumulated effects of stress which lead to strengthenings

⁴¹ Co-occurrence of vowel length and stress is a widespread phenomenon, and can be also found e.g. in Slovene (Šuštaršič & Komar 1999) or Italian (D'Imperio & Rosenthal 1999).

⁴² However, Jacewicz et al. (2006: 312) state that this strengthening can manifest in dialect specific ways, i.e. whether the strengthenings are taken up depends on "the social soil and climate in which this seed is planted", as they show that /eɪ/ raises, i.e. strengthens, in south-central Wisconsin, but not in central Ohio.

would be to assume that phonemes which are frequently associated in production and perception with strong prosodic positions, begin to be stably linked to those positions, that is acquire characteristics typical of those positions. This is plausible under current views of Cognitive Grammar (Bybee 2001: 20, citing Langacker 1987), which assumes that the presence of a regularity does not have to mean that the information predictable from this regularity is absent from the individual forms that are subject to it, and it is also very much at home in the association network scenario of evolutionary linguistics (Ritt 2004), where it is assumed that the links between neural patterns which are frequently activated together become ever stronger, so that it becomes ever more possible that the excitation of one of them would activate the other⁴³.

For example, the stressed vowel of OE *mētan* ‘meet’ was consistently phonetically long. It was a phonologically long vowel and so it had inherent length. Further, this inherent length was rarely phonetically shortened, since the stress was not shifted to any suffix in other forms, and stress-induced phonetic effects on the duration of the vowel, if any, were bound to be those of lengthening. The long-term result of that would be strengthening. If the phoneme /e:/ found in words such as *mētan* was activated frequently with the prosodic structure |óσ|, to yield forms such as /'me:tan/, then an easily excitable link came to be established directly between /e:/ and |ó|. As a result, each activation of the phoneme would also activate the link to a strong prosodic position, and each production of the phoneme would display the characteristics typical of that prosodically strong position.

Under such a model, the strengthening of a segment would be linked to the consistency with which it is stressed. It would allow for the concentrated effects of language use to be manifested. It can be speculated that just as each lexical entry is equipped with a frequency metric⁴⁴ (as suggested by Pierrehumbert 2006), so is each phonemic position equipped with another frequency metric, namely one counting the number of times a given phoneme found itself in a prosodically strong position, that is in a foot head. The value of this metric would not be determined by the composition of a given word alone, but would be influenced by the

⁴³ There would be even more incentive to store, and draw on, those permanent links once the morphological conditions for stress assignment have become less transparent in Early Middle English (Ritt 2012: 400).

⁴⁴ which accounts for lexical frequency dependent reductive changes

words with which a word co-occurs in speech, since even a sound in a foot head of a lexical item can be demoted or lifted in actual utterances. The value of this metric would then determine the degree of hyper- or hypoarticulation of the sound, which accounts both for the strengthening effects seen in frequently stressed syllables, and for weakening effects seen in frequently unstressed syllables. Just as the frequency metric for entire lexical items gets updated through language use, so is the ‘segmental strength’ metric updated with each association of a segment with a strong prosodic position. A particular segment which is consistently stressed and therefore repeatedly given prominence enhances these characteristics and becomes even more prominent with time. Taken together, the underlying phoneme, the frequency of the lexical item it is embedded in and the stored frequency with which it has been associated with a strong prosodic position allow the computation of the actual phonetic form.

To allow for the continued accumulation of the effects of language use on grammar, one can employ the model outlined above, with the actual phonetic form depending in part on the two frequency metrics, the one of the lexical item and the one of the segmental strength. On this view, however, the representations of individual phonemes remain unchanged, at least with regard to the reduction/strengthening scale, it is only the metrics, and with them the phonetic realizations that change. To account for historical change of phoneme lineages over time, then, one has to assume that an important role is played by language acquisition. Vowel shifts are clearly non-reductive, across-the board changes, so they are not driven by high frequency of the words that the vowels are in. If one sees them as strengthenings due to the frequent association with a strong prosodic position as outlined above, though, the same mechanism would apply to them as well. The strengthening gets entrenched in the lifetimes of individuals, and becomes a permanent characteristic of their speech production, but the change in the representation of a category occurs through the transmission of the phoneme to a new brain, where a new phonemic category is established based on the characteristics of the output forms available in ambient language.

In this way, in a language with fixed lexical stress, stressed syllables can be conceived to accrue the phonetic effects of stress. In a word such as *rīden* ‘ride’, for example, even before the loss of the second vowel, it is the first one that was invariably stressed, and so it was invariably expressed with a long phonetic duration, which allowed it to strengthen, or to

tense. After final schwa loss, as such words became monosyllables, the likelihood of strengthenings was under two contradictory pressures. On the one hand, since the stressed *rī-* of *rīden* could not rely on being always followed by an unstressed syllable, it ran a risk of clashing with another stressed syllable and therefore of rhythmic demotion. On the other hand, the reason why strengthenings were more likely is that not even under special circumstances could the prominence relation between the first and the second vowel be more even. Ultimately, another strengthening, namely diphthongization, followed, suggesting that the combined influence of low morphological functionality of final schwas and the growth of the difference in prominence between stressed and unstressed syllables were more decisive than the need to avoid stress clashes (for a suggestion regarding the mechanism behind schwa loss, see Section 4.3).

The segmental effects of stress are expected to be different in languages with mobile stress, that is stress which can fall on different syllables within a paradigm. This is for instance the case in East Slavic languages (i.e. in Russian, Ukrainian and Belarusian, which have word stress which is both free, as in the Greek example above, and mobile, see Sussex & Cubberley 2006: 15) and in Romance languages, including French and Italian. In Italian, for example, as a result of the existence of stress attracting inflexional suffixes, stress can fall not only within a stem but it can also be shifted onto the suffix. In Italian verbal morphology, for instance, inflectional suffixes can affect stress placement by shifting it away from the stressed vowel of the stem to the final syllable (e.g. future) or to the penultimate syllable (e.g. imperfect indicative, subjunctive) of the complex form, as illustrated below with the forms of the word *parlare* ‘to speak’ (examples adapted from D’Imperio & Rosenthal 1999: 21).

(3) *a. initial stress (on the root)*

párlo	párli	párla
speak.1SG.PRS	speak.2SG.PRS	speak.3SG.PRS

b. final stress (on the suffix)

parleró	parlerái	parlerá
speak.1SG.FUT	speak.2SG.FUT	speak.3SG.FUT

c. penultimate stress (on the suffix)

parlávo	parlávi	parláva
speak.1SG.IPFV.IND	speak.2SG.IPFV.IND	speak.3SG.IPFV.IND
parlássi	parlássi	parlásse
speak.1SG.IPFV.SBJV	speak.2SG.IPFV.SBJV	speak.3SG.IPFV.SBJV

As a result of this movability of stress, it is expected that in such a language the stressed vowels will accumulate the strengthening effects of stress to a lesser degree than in languages with more stable stress. Admittedly, segmental effects of stress in the synchronic phonology of Italian are observable, namely there is phonological lengthening of stressed penultimate open vowels, and phonetic lengthening of stressed antepenultimates (D'Imperio & Rosenthal 1999), but no further strengthening seems to be taking place, presumably due to the fact that stress is not consistently linked to the same syllable for a given lexical entry. Needless to say, the segmental effects of stress are expected to be even less pronounced in languages with no lexical stress, such as Japanese for example.

In contrast to languages with movable stress or languages with no lexical stress, where the cumulative effects of stress are expected to be slight, they are expected to be much more extensive in English. In English, over time, the difference in prominence between stressed and unstressed syllables has become greater. As a result, since a trochee is a sequence of a strong and a weak syllable, many disyllables can be seen as now yielding better trochees, as the strong syllable became even stronger, and the weak syllable became even weaker. And so, fixed lexical stress and the concentrated segmental of it contributed to an enhanced expression of trochaic rhythm, at least until the time of schwa loss, when an increased number of monosyllables would actually interfere with it.

This development, where vowels in unstressed syllables were progressively weakened and vowels in stressed syllables progressively strengthened, has brought OE closer to the ideal stress-timing language type. The argument that the concentrated segmental effects of stress result in strengthenings and therefore possibly in shifts is in agreement with the observation that shifts affect stressed vowels only (Donegan 1993). It alone points to a link between stress and shifts. This argument ostensibly seems to be challenged, though, by Bybee et al.'s (1998) study, which found a negative correlation between the predictability of stress placement and the segmental effects of stress. These results at first seem to contradict the mechanism sketched here, where the fixed stress of OE is argued to have brought about the accumulation of segmental effects of stress. The fixed stress of OE was at the same time very predictable, and Bybee et al.'s (1998) study suggests that predictability of stress placement reduces its segmental effects. This finding, however, is challenging only at first blush. It points, in fact, to the possible rupture between the effects of predictability of stress placement on the one hand, and of 'fixedness' or lexical stability of stress on the other. While the segmental effects of stress are able to accumulate only when stress is fixed on one and the same syllable, as outlined above, there is not a great incentive for this to happen when the stress is very predictable, that is grammatical. Thus, the very fixed and relatively transparent stress assignment of Germanic, continued in Old English, had a seed of the accumulation of segmental effects of stress (through lexical fixedness), but it at the same time carried with it a countermeasure (a high degree of predictability). When stress remained tied to particular syllables through ME, and therefore 'fixed', however, while conditions for stress assignment became much more complex than the simple leftmost stress of OE (due to the obscuration of morphological transparency, discussed further in 4.3.3 below), and therefore stress placement was less predictable, the segmental effects of stress started to accumulate.

To recapitulate, fixed lexical stress in OE enabled the accumulation of segmental effects of stress, which (a) enabled the strengthening of stressed syllables and (b) lead to the weakening and loss of unstressed syllables. The growing contrast between stressed and unstressed syllables, in itself a feature conducive to stress-timing, contributed to the loss of the unstressed vowels, which stands in the way of stress-timing through the increase in the number of stress clashes. Although the growing dominance of trochaic rhythm should have prevented schwa loss in contexts where schwa loss endangered rhythmic alternation, that is

when it led to the shortening of disyllables to monosyllables, it did not do so and schwa loss proceeded. As recounted in 3.3.1 above, low functionality of highly syncretic inflections contributed to schwa loss. As a consequence, other strategies to enhance rhythmic alternation developed, which is the topic of Section 4.3.

4.2.3 Fixed lexical stress as a factor undermining the stress-timing of English though causing stress clashes

The second way in which fixed lexical stress is important for the weakening of the reliability of duration as a marker of phonological length contrasts is, somewhat paradoxically, in another kind of interaction that it enters into with the rhythm of English, an interaction which undermines, instead of enhancing, the stress-timing nature of the language. Namely, if one and the same syllable of a word is stressed, then for each syllable of the word the probability that it will find itself in a sequence characterized by a stress clash or a stress lapse becomes greater than it would be the case if the stress were movable.

To refer to the Italian example of a paradigm displaying mobile stress presented in (3) above on page 150, it can be noted that because the first syllable of *parlare* is stressed in some inflected forms but not in others, it is subjected to fewer stress clashes than it otherwise would be. Namely, if the first syllable were invariably stressed, then for each utterance in which *parlare* directly follows a stressed syllable, a stress clash would ensue. However, since the first syllable is stressed only in certain word forms, then a stress clash will ensue only in those cases of *parlare* directly following a stressed syllable, in which the stem is stressed. Hence, when stress becomes tied to a given syllable in a word, the likelihood of this syllable occurring in a stress clashing context increases (unless, of course, the lexicon of the language in question were composed exclusively of initially stressed disyllables).

Thus, there is a complex relation between fixed lexical stress and a tendency for stress timing. Auer (1993: 94), for example, includes ‘word accent’ as a prototypical characteristic feature of ‘word based’, i.e. stress-timed languages. This is only logical since word prosody is linked to strong stress, and strong stress is a characteristic feature of languages with fixed word accent. But on the other hand, fixed lexical stress leads to stress clashes when a number of stressed monosyllables are thrown into the mix. In an idealized scenario where OE

consisted only of words of the shape (óσ), their sequences would always yield perfect binary rhythm. If this idealized scenario of initially stressed disyllables is disturbed by an introduction of stressed monosyllables, however, then sequences with stress clashes, such as (ó)(óσ) or (ó)(ó) surface. Each individual lexical item will be subjected to such clashes more frequently than it would be if it had more movable stress, since, with a movable stress, some forms of the word would have the stress on a different syllable. It also leads to sequences of unstressed syllables, or lapses. Now, there is a phonetic motivation to alternate weak and strong syllables. When the effects of this phonetic motivation, namely phonetic weakening of one of the stressed syllables to assuage a stress clash, or phonetic strengthening one of the unstressed syllables to assuage a stress lapse become reinterpreted as intended and thus become phonologized as part of the phonetic component, a process of phonological demoting of stressed syllables and promoting of unstressed syllables might arise. Both clashes and lapses stand in contradiction of the preference for alternating stress, and two possible strategies to resolve this unwanted consequence are to either promote unstressed syllables/demote stressed syllables prosodically or to stretch/shrink syllable durations phonetically. This effect of fixed lexical stress, namely the increased rate of stress clashes and lapses, must have been operative already in OE. As for the phonetic and/or phonological resolution of stress clashes and lapses, there is clear evidence for it for ME, as illustrated in Ritt (2012), to be discussed in Section 4.3 below. The evidence of the prosodic strategy being employed in Middle English (Ritt 2012) points to the robustness of the tendency for alternating stress at that stage. What is now in order is the discussion of the status of the aforementioned ‘strategies’.

As is clear from the discussion in Section 3.3.2, linguistics has yet to establish the feasibility of a rhythm-based typology, let alone work out its exact details. Still, every attempt at approaching such a typology has placed English close to or on one of the extremes of the proposed scale. Dauer’s (1983) study combining a number of phonetic and phonological traits found English to be the most ‘stress-based’ of the languages investigated. Auer’s (1993) attempt at a prosodic typology, though it concludes by claiming that such a typology should not be based on rhythm but rather on the prosodic domain central for a language’s phonology, places English at one of the extremes of the scale, equaling it with a

prototype ‘word language’. Grabe and Low’s (2002) phonetic investigation yielded high consonantal and vocalic PVI values, pointing to English being more ‘stress-timed’ than numerous other languages in the study. Especially the approaches arguing for a scalar typology and incorporating phonology, such as Dauer (1983), or even limited to phonology, as Auer (1993), are compatible with a historical interpretation. Throughout its history, English has moved ever closer to the word language prototype by its phonology meeting ever more criteria of being a typical word language. One feature typical of Auer’s prototypic word languages, for example, is the occurrence of a restricted number of vowels in unstressed position, and English acquired this trait by the eleventh century, with all weak vowels except /i/ merging into one (Lass 1992a: 77). Even though it is not yet understood if and how the criteria might influence each other’s development, it seems safe to assume that they do, and such an assumption is particularly plausible from the evolutionary perspective. As one replicator, namely the trochaic foot⁴⁵ propagates, and is successful in doing so since, as numerous empirical studies suggest⁴⁶, the rhythmic characteristics of a language are acquired very early on in the language acquisition process, those other replicators will succeed which replicate well in an environment which contains the said trochaic foot. So for example, lexical entries which have only schwas as their unstressed vowels, replicate better if a trochaic foot is their co-replicator, than words with full vowels in the unstressed position would, since the weak/strong opposition between the head and the non-head of a trochee is more pronounced if the non-head contains a reduced vowel. By the same token, if the vowel sitting in the head of a trochee raises, so that it strengthens (Donegan 1978), it makes the trochee it is part of better as trochee. However, the surfacing of a reduced vowel in the non-head, or the raising of the vowel in the head cannot simply fall out of the constraint ranking relevant for the tendency for the trochaic foot, namely highly ranked FTBIN and TROCHAIC outranking IAMBIC (or the presence of TROCHAIC rather than IAMBIC). This raises, for one thing, the question about the nature of the interplay between the preference for binary trochaic rhythm as a feature of English prosody which has been more fully realized throughout its history, and the synchronically operative ‘strategies’ remedying the surfacing of utterances violating this rhythm.

⁴⁵ in OT, at the word level, a particular ranking of FTBIN, PARSESYL, TROCHAIC, and possibly other constraints, cf. Kager (1999: Chapter 4) or Apoussidou (2006: Chapter 2)

⁴⁶ E.g. Mehler et al. (1988); Dehaene-Lambertz & Houston (1998); Nazzi, Bertoncini & Mehler (1998)

The relevant constraint ranking was there before the (historical) vowel reduction and before the (historical) vowel raisings, it was there while they were happening and it has been there afterward. Consequently, the diachronic developments of the optimization of the trochee (or the historical ‘tendency’ for the trochee) are not an automatic result of a particular constraint ranking. However, in discussions of synchronic phonological systems the term ‘tendency’ is used to refer to just that, i.e. to particular constraint rankings.⁴⁷ In other words, a ‘tendency’ for a certain kind of rhythm in OT is the state of affairs at a given synchronic stage of the language, which is essentially the optimality of certain kinds of prosodic structures as against the sub-optimality of others. This ‘tendency’ is simply an effect of a particular constraint ranking, and it could even be equated with the foot itself in such usage, because a foot is not given an ontological status separate from the relative constraint ranking. Thus formulated, the presence of a tendency seems to have nothing to say about diachronic developments.

In contrast, the preference for the trochaic foot, as it is understood here, is not identical with a constraint ranking which yields trochaic feet. It includes it, since the presence of that constraint ranking is the result of phonologization of the phonetically, or, ultimately, physiologically motivated preference for binary alternation, and for a single kind of alternation rather than a mix. However, claiming that there is a preference for trochaic feet, manifested in long-term developments of a language, cannot be limited to stating that there is such a phonologized component. If a preference is to be historically operative, it must mean that there is a certain readiness on the part of the language to keep incorporating the physiologically motivated preference into its grammar. This readiness could result from the existence of lexical items which realize the trochaic foot, as they provide an environment in which the replication of new competence constituents which comply with it is facilitated. This is slightly different to the, somewhat similarly puzzling, invasion of the grammar by rules of increasing generality referred to by Bermúdez-Otero (2013 (forthcoming)). There,

⁴⁷ In Kager’s (1999) usage, for example, a ‘tendency’ for trochaic feet would have to be split in (at least) two. First, a tendency for rhythmic alternation is just the surfacing of binary feet. It is the result of high ranking of FTBIN and PARSESYL in a particular language. The tendency for those binary feet to be binary results from RHYTHMTYPE=TROCHAIC outranking RHYTHMTYPE=IAMBIC. At any rate, even assuming that conflating all those into one makes sense, a tendency for trochaic feet is shorthand for a description of a synchronic generalization that surface forms of a given language are characterized by high incidence of binary trochaic feet.

processes of increased generality enter the grammar from phonetics through successive higher levels of the hierarchy, which refer to similar, but increasingly more general, domains. Here, processes referring to different domains (such as an increase in qualitative differences between vowels or consonantal simplifications) keep entering the grammar for the benefit of an extant feature of the grammar, namely the trochaic foot.

At any rate, even if the typological shift toward stress-timing is thought to have become manifested in the phonological domain, the processes involved must have begun their life-cycle in the phonetic domain, with changes in syllable complexity necessarily ultimately deriving from changes in the duration of vocalic and consonantal intervals. This, to return to the strategies of repairing stress clashes and lapses, is of consequence for the meaning of the way in which the ‘tendency’ for the trochaic foot makes its way into grammar. As the origin of all sound patterns derives ultimately from phonetics, it must be assumed that such a prevalent feature of spoken languages as alternating rhythm (cf. Dziubalska-Kołaczyk 2002: 89 and references therein for a number of arguments for the universality of binary feet, and of the trochee specifically) has some physical basis. The general limitations of signal processing in humans, captured by the semiotic figure-and-ground principle, and specifically the dislike of consecutive figures are argued by Dziubalska-Kołaczyk (2002: 91) to be the grounding of a preference for binary foot structure. One proposal suggesting that the preference for binarity might have its motivation in the mechanism of speech production is the theory that the whole of phonology is an exaptation of the cyclical properties of the mandible (MacNeilage 1998). As such, a ‘tendency’ is a characteristic of the physiology of speech production, which results in the typological frequency of phonologies reflecting this characteristic. Thus, the pressures of speech processing and speech production converge on the recurrence of binary feet, the phonologization of which, in OT terms, is a highly ranked FTBIN constraint.

If the suggestion that the existence of a synchronic ‘tendency’ for the trochaic foot is the result of phonologization, stabilization and domain narrowing of the physiological tendency for binary structure, then the existence of such a synchronic ‘tendency’ (in the sense of a particular constraint ranking, yielding trochaic feet at the word level, in the sense of Kager 1999: Chapter 4) must be seen as resulting of a previous life-cycle of this kind (cf. Figure 3 on on page 74). The fact that this life-cycle has proceeded to the word level already in the

past does not in any way block its reappearing at the post-lexical level once again. The existence of a phonologized trochaic preference on the word level does not preclude a physiologically driven phonetic regularity creeping into phonetics through phonologization once again. Such a duplication can be seen as an example of ‘rule scattering’ (Bermúdez-Otero 2013 (forthcoming)). The physiological motivation for it is simply still there, and so it can creep into phonology at the post-lexical level, even though it had already made it once all the way to the word level.⁴⁸

Bringing all those perspectives on the relationship between ‘preference’ and ‘tendency’ together, a preference for trochaic binary feet is the presence of such feet, or the relevant constraint ranking, in the grammar, together with the effect that the presence of this feature has on diachrony, which is accepting such innovations into the phonetic component which enhance it. Synchronically, it is visible in the grammar as a process operative at different levels, with the ultimate motivation being in the physiology of speech. Diachronically, the preference for the trochaic foot is the totality of changes which make surface forms realize the trochaic foot better.

4.3 MIDDLE ENGLISH

Middle English saw the continuation of the accumulation of segmental effects of stress, a trend which began already with the fixing of lexical stress in Germanic as argued above. It also clearly saw the application of the strategies for avoiding stress clashes and lapses (possibly operative already in OE), whose urgency had become greater than before as numerous unstressed vowels were reduced and then lost.

Now, as far as the segmental effects of stress are concerned, the process of the erosion of vowels in unstressed position, which began in OE, progressed through ME. While in OE length distinctions were leveled out, in ME vowel quality distinctions were lost as well, with most vowels in unstressed position merging in schwa by the eleventh century (Lass 1992a: 77), and with final unstressed schwas being lost altogether by about 1400 (Lass 1992a: 79).

⁴⁸Alternatively, the binary structure requirement on the post-lexical level is not an innovation, but a remnant of a previous life-cycle. These two cases are equally plausibly, and would be, from a synchronic point of view, indistinguishable.

These developments were not without influence on the way in which rhythm interplayed with segments. It could be speculated that the increased incidence of stress clashes and lapses, which is concomitant with the fixing of stress, was remedied in production by speakers already in OE. While it remains not much more than a speculation for OE, there is evidence of this happening in ME. In OE, and through ME until complete schwa loss, stress lapses, if any measures were taken to militate against them, could hypothetically be remedied by promoting or stretching vowels in unstressed syllables, and stress clashes could be remedied by the optional retention of a final schwa. And in fact, schwas were retained the longest, or even occasionally inserted, in weak singular adjectives preceding nouns, which can be seen in the examples given by Minkova (1991: 181), recounted below.

- (4) a. The newe shame of Sennes olde (Gower, *Confessio Amantis*, (vii, 5115-6), ed Macaulay (1900 & 1969)
 b. a riche feste (Gower, *CA*: ii.4702)

Example (4).a above shows a retention of morphological schwa attached to a monosyllabic adjective, which yields rhythmic alternation. Here, it agrees with the old grammatical motivation, where *-e* is expected in the definite context. In 3.b., however, there is no grammatical motivation, since the context is indefinite. The presence of examples like this one is adduced by Minkova (1991: 171) as evidence for rhythmic, rather than grammatical, causes behind the lingering retention of adjectival schwas, and so it can be thought of as a kind of strategy applied by speakers to maintain alternating rhythm.

It is conceivable that such strategies would be preferable to strategies involving stressed syllables, since stressed syllables carry “a maximum of information” (Auer 1993: 6), and any manipulation of the input for rhythmic purposes which obscures the identity of a segment is more tolerable if it is a segment which carries less information. Hence, it is conceivable that the relevant adjustments should have affected vowels in unstressed syllables before they affected vowels in stressed syllables. The strategy of promoting unstressed syllables became much more restricted, however, and the strategy of the optional retention of the unstressed syllables which were normally deleted became unavailable in many cases in late ME, as many unstressed vowels were lost altogether by the early fifteenth century (Lass 1992a: 81).

Given that final schwas were useful in maintaining alternating rhythm, it is more than a little puzzling that they were lost. To address this bewilderment, one could try and treat final schwa loss as a case of the so-called evolutionary suicide, a type of evolutionary extinction which is an evolutionary development in which successive adaptations, though each of them is beneficial to the individuals concerned, ultimately lead to the extinction of the population, or where “the population actively evolves toward self-destruction, i.e., mutants closer to the extinction boundary are at advantage with respect to resident individuals, even though they are unconsciously closer to extinction” (Dercole & Rinaldi 2008: 23). A likely example of such an event in biology is the extinction of the family *Brontotheriidae* of the order *Perissodactyla*. These mammals, similar to their modern day relatives, namely horses, differed from them by evolving horns and large body sizes. Though the cause of their extinction is still debated, it is very likely that their own large bodies led to their demise (Dercole & Rinaldi 2008: 28, 30).⁴⁹

Schwa loss could be seen from this perspective along the following lines. The progressive weakening and eventual loss of final schwas in disyllabic words can be schematically represented as follows:

- (5) a. SW e.g. GMC *stainōs ‘stone.NOM.PL’
- b. Sw e.g. OE /ˈstaːnas/
- c. S_w e.g. ME /ˈstɔːnəs/
- d. S e.g. PDE /stəʊnz/

The stage represented at (5).a above corresponds to the situation in GMC, when the unstressed syllables in initially stressed disyllables could contain a full range of vowels. Stage (5)b. corresponds to the situation in OE, where quantitative contrasts in weak syllables were leveled out, and stage (5)c. to a situation in late OE, where the number of qualitative contrasts in weak position was limited. (5)d. represents a stage in late ME, when many schwas in unstressed position were lost altogether. In this progression, one could argue that

⁴⁹ A less extreme case of an adaptation potentially harmful to the population is the peacock’s tail (Parvinen 2005: 241). Although it is beneficial to the individual, by raising its chances of mating, it is injurious to the population in the long run, as such a conspicuous appendage raises the list of the peacock falling prey to predators.

b. is a better trochee than a., and that c. is a better trochee than b. Since a trochee is a binary foot with the first syllable more prominent than the second, this succession represents trochees of increasingly large difference in prominence, and in that way a succession of better trochees. For a development to be a case of evolutionary suicide, two conditions have to be met (Parvinen 2005). First, it must be the case that for each successive stage, the new variant is able to replace the old variant. This is clearly the case for stages from a. through c. Words with no quantitative contrasts in their weak syllables took over and replaced words which had those contrasts, and later words with few qualitative contrasts in their weak syllables took over and replaced words which had such contrasts. Second, it must be the case that the last innovation in this suicidal chain manages to replace its predecessor, but that it is not in itself stable, leading to its dying out⁵⁰. This last step is the most interesting one in case of schwa loss. While the succession from a. through c. is plausibly seen as structures producing better forms from a rhythmic perspective replacing worse structures, this is not the case with the last innovative form. A one-syllable word is, on its own, not able to produce a binary foot. The advantages that it could have had in its favor, though, are the following. First, it is shorter, and so requires less effort, a factor that gained in importance throughout the process of weakening of unstressed vowels, because it was no longer offset by morphological functionality. Second, with no other syllable present, a stressed syllable in a one-syllable form does not run the risk of losing its prominence to another syllable through stress adjustment for rhythmic purposes, and so a one-syllable word has a guaranteed greater invariance with regard to stress placement. These benefits must have initially outweighed the benefits of retaining a schwa, which yielded binary feet regardless of what kind of a word followed the word in question. Their rhythmic inadequacy, however, eventually led to the demise of one-syllable words which employed syllables as their unit of quantity, as evidenced by the rise of moraic feet in English. Hence, FTBIN started to operate with regard to moras, rather than syllables, at least on the word level. This short sketch does not solve the issue of why final schwas were lost, but, seeing this development as a case of evolutionary

⁵⁰ One might posit an analogy from the plant world to illustrate this somewhat counterintuitive development. For example, a mutant tree, taller than the residents, is able to gain more sunlight and therefore replaces the resident. However, its very size means that it requires more natural resources than available, such as water, and as a consequence the new population dies out by depleting their environment (Baumann 2012).

suicide points to the kind of evolutionary scenario that might help uncover the conditions under which it unfolded.

At any rate, numerous final schwas were in fact lost, and so these were no longer available, neither for promotion nor for optional retention. With these strategies gone, stressed vowels, which could previously be spared from manipulation, became subjected to demoting and stretching/shrinking (as illustrated in 4.3.2 below). E.g. when two lexical monosyllables clashed, one would lose its prominence post-lexically to yield binary rhythm. Thus, the duration of vowels in stressed syllables came to be co-determined, on top of their phonemic length and stress, also by the presence of other stressed syllables in their vicinity.

There is an additional way in which the loss of final schwas (which was itself precipitated by the fixing of stress and progressing syllable-timing) interplayed with rhythm, and which contributed to the rise in variability of vowel duration. Not only did the loss of final schwas (when it was complete, i.e. most likely by around 1400 Lass 1992a: 79) make the duration of stressed vowels more variable by decreasing the availability of unstressed vowels which could be promoted or optionally retained, but it also increased the number of cases where a strategy to achieve rhythmic alternation was needed in the first place. It shortened a substantial number of words, resulting in a vast increase in the number of monosyllables. As a consequence, occurrences of stress clashes, that is of consecutive stressed syllables, became more frequent, and the incidence of stress clashes, more prevalent with fixed stress than with mobile stress, and frequently undermining rhythmic alternation on the post-lexical level, which must have been present already in OE and early ME, now rose even more. This, together with a decreased availability of unstressed vowels to achieve rhythmic alternation, has led to a situation where the frequent clashes must have resulted in utterances with no binary rhythm, that is in cases where post-lexical adjustments were needed.

As a result of the duration of stressed vowels being influenced not only by their phonemic identity and by stress, but also by the post-lexical adjustments, the duration of stressed vowels must have become highly variable. It is suggested here that at that point the inherent qualitative differences between short and long counterparts of the phonemic oppositions (whose existence was argued for in Section 3.4) were co-opted as primary

markers of the contrasts. This exaptation of the inherent qualitative differences is presented in greater detail in Section 4.4 below.

The application of the strategies which led to the increase in variability of vowel duration is made explicit below by invoking the relevant constraints and constraint rankings. It is tightly-knit with the re-organized stress assignment in ME, which has to be presented first.

4.3.1 *Stress assignment in ME – grammatical stress*

First of all, it has to be noted that stress assignment on the lexical level has become weight-sensitive (1992a: 85ff.). It is a widespread generalization that stress tends to coincide with heavy syllables, and in pre-OT phonology it was variously referred to as “Prokosch’s Law” (Prokosch 1939) or Weight Law (Vennemann 1988) (both cited by Minkova (2006: 115). Within OT, this generalization is expressed as the presence of a markedness STRESS-TO-WEIGHT constraint.

(6) STRESS-TO-WEIGHT (STW): If stressed, then heavy⁵¹

The introduction of weight-sensitivity into English phonology is usually linked to the extensive influx of French and Latin loanwords into English in ME. However, the co-occurrence of weight and stress was present already in OE. First, lexical monosyllables were all heavy already in OE. Second, in polysyllables, a heavy syllable was usually stressed. These two facts paved the way for weight-sensitivity in ME. This co-occurrence got tightened with the Middle English Open Syllable Lengthening (MEOSL), when a number of short stressed vowels were lengthened, resulting in more cases of correlation between weight and stress.

The rise of weight-sensitivity was interpreted as the introduction of the ‘Romance Stress Rule’ (Halle & Keyser 1971: 100) in accounts couched in rule based generative phonology. In OT, no such rule is needed, and the generalization expressed by it is captured instead by the interaction of STW with other constraints. Minkova (2006) suggests that the constraints

⁵¹Note that this constraint (which Kager [1999: 268] traces back to Myers [1987]) penalizes forms in which foot heads are not heavy. The cross-linguistic co-occurrence of stress and heavy syllables, if one wants to express it with markedness constraints, is not fully captured by STRESS-TO-WEIGHT, though. To do so, it has to be complemented by another constraint, namely by one penalizing forms in which heavy syllables are not foot heads. This complementary constraint is the WEIGHT-TO-STRESS PRINCIPLE (WSP), formulated by Prince and Smolensky (1993 [2002]: 56), who trace it back to Prince (1990).

needed to do so are the already mentioned STW and FTBIN, as well as, not mentioned so far, NONFINALITY. It is a constraint capturing the “avoidance of stress on final syllables” which is “a very commonly encountered phenomenon in stress systems of all kinds, typically attributed to various forms of extrametricality, stress-shift, and de-stressing” (Prince & Smolensky 1993 [2002]: 41).

- (7) NONFINALITY: The prosodic head of the word does not fall on the word-final syllable (Prince & Smolensky 1993 [2002]: 42)

Weight-sensitivity of stress assignment in ME suggests that feet were constructed out of moras now. This means that FTBIN, at play in ME as much as it was in OE, must have changed the domain of its application, and must have operated with regard to moras in ME, rather than syllables, as it used to be the case in OE. Lexical monosyllables were in the minority in OE, and it makes sense to assume that their rarity is matched by the exceptional mechanism of catalexis. In ME, however, monosyllables became much more abundant, at least in the native vocabulary. As a consequence, it seems to be no longer feasible to invoke an exceptional status to account for the forms which are now highly frequent. Assuming that FTBIN referred to moras and not to syllables in ME is a solution to this quandary. Lexical monosyllables are binary on the moraic level, since they are all heavy syllables.

4.3.2 Avoidance of stress clashes and lapses

In view of the above, mora-oriented weight sensitive grammatical stress assignment can be postulated to have functioned on the lexical level in ME. As the cases of stressed syllables being demoted in Middle English suggest, however, phonetically driven reduction of post-stress syllables must have been, by that time, phonologized, leading to adjustments of stress post-lexically. To represent this situation in OT, there are two possibilities. Either the metrical constraints apply to a domain different than a word, which is unlikely given the extensive body of research constructed around the word as a metrical domain; or the optimal candidates of the word level are inputs to another level, say a post-lexical level. This second option goes against the tenets of classical OT but could be accommodated within Stratal OT.

Assuming that post-lexical stress adjustments were brought about by the submission of the output of the lexical level to the post-lexical level, the following account shows how the

strategies for stress adjustments worked. No constraints exclusive to the post-lexical level have to be posited. What is posited, though, is that two constraints, namely FTBIN and PARSE, have differing requirements, depending on the cycle at which they apply. The following are examples of post-lexical stress demotion and promotion yielding rhythmic alternation on the post-lexical level. These are examples considered by Ritt (2012: 404), and his proposals for stress assignment are modeled here. The first one is an example of a lexically stressed syllable, namely the only syllable of the word *right*, which loses its stress on the post-lexical level.

- (8) For, 'quyk or 'deed, **right** 'there ye 'shal me 'fynde. (c.1390 Chaucer, “Franklin’s Tale”, CT F628)

Tableau 5 below shows the application of the relevant constraints. For the simple scenario presented therein, it suffices to employ the constraints responsible for stress assignment on the lexical level in OE. What would have to be assumed, though, is that even though for lexical stress assignment FTBIN has switched to the mora as its basic unit, it still operates with regard to syllables in the post-lexical cycle.

‘deed right’ Lexical level: (ó)(ó)	FTBIN	RHTYPE=TROCHAIC
(ó)(ó)	*!	
íð (óσ)		
(σó)		*!

Tableau 5: Demotion of stress on the post-lexical level

An alternative to assuming that FTBIN has two domains depending on the cycle would be to invoke a separate constraint, namely *CLASH, a markedness constraint against sequences of stressed syllables. It was posited by Minkova (2006) to account for stress assignment on the lexical level, so again no phrase-specific constraint would have to be posited. In defense of using *CLASH, it can be said that it seems realistic that only one kind of feet is chosen in a language, irrespective of the cycle. On the other hand, if FTBIN could do the job, no specific clash avoidance constraint would be needed, at least on the post-lexical level.

The second example is a case of a lexically unstressed syllable, namely the preposition *in*, gaining prominence on the post-lexical level.

- (9) ne 'suffre þu 'nawt þe 'feond þt ne 'lead(e) us 'allung(e) ,**in** to 'fondung(e) (c. 1230 Ancrene Wisse: 116)

There are two ways in which promotions could be accounted for. One is to posit that an additional constraint, such as *LAPSE employed by Minkova (2006) is at play, another, the one employed here as exemplified in Tableau 6 below, is to assume that the footing requirements on the post-lexical level are stricter than on the lexical level, namely all forms are subject to PARSE on the post-lexical level. PARSE is a rhythmic constraint which requires that all syllables are part of feet. It has not been employed so far, as it would not have played a role in the stress assignment cases on the lexical level presented so far.

- (10) PARSE: Syllables are parsed by feet (Kager 1999: 153)

In other words, while grammatical words such as prepositions can remain unfooted lexically, they are not exempt from rhythmic constraints on the post-lexical level.

'allung(e) in to' Lexical level: (óσ) σσ	FTBIN	RHTYPE=TROCHAIC	PARSE
(óσ) σσ			**!
☞ (óσ) (òσ)			

Tableau 6: Assignment of secondary prominence on the post-lexical level

The existence of such examples strongly suggests that, especially once the final unstressed vowels were gone, then, stressed syllables could be demoted to avoid stress clashes, and unstressed syllables could be promoted to avoid stress lapses. It has been suggested before that clash and lapse avoidance was operative on the word level in ME (Minkova 2006), and examples like these suggest that it was also active post-lexically. Given the view on the relationship between word-level and utterance-level phonology formalized in the life-cycle of phonological process, not only is such a co-existence plausible, but it would in fact be expected that the word-level process originated as a post-lexical one.

4.3.3 Stress assignment in ME – lexical stress

Stress assignment in ME can in general be modeled with the use of metrical constraints, as depicted in 4.3.1 above. However, there are reasons to believe that stress in ME began to, at

least occasionally, be marked lexically. First, there are lexical items that behave exceptionally and for which stress cannot be predicted from their phonological and morphological composition. For example, forms such as *a'būte(n)* ‘about’, *ʒe'nōh* ‘enough’, *a'ʒeines* ‘against’, *a'wei* ‘away’, *bi'leafe* ‘believe’, *bi'hālden* ‘behold’ were stressed on a syllable other than the first syllable of the root, since the morphological composition of ‘prefix + root’ was no longer transparent (Ritt 2012: 400). The loss of morphological compositionality in these forms did not result in stress being placed on the first or the leftmost syllable of the (new) roots, though. To account for that, one has to assume that these became lexical items with lexically marked stress. As the morphological transparency of certain forms was lost, so was the grammatical conditioning of stress, and lexical marking of stress became an option in English, at least in the limited number of such examples. Second, departing from the clear-cut separation of information being stored either in the lexicon or in the grammar, it is assumed here that stress could be marked lexically even in items in which it was still predictable.⁵²

In OT, allowing stress to be marked directly in lexical entries does not suffice to make sure that it will be manifested in the output. One also has to posit relevant constraints which will enable stress markings to survive the journey from the input to output forms. One way of doing so is to posit, in the spirit of Correspondence Theory (McCarthy & Prince 1995a), MAX and DEP constraints, which militate against deletion and addition of features between the input and the output. This is done by Revithiadou (1998: 26), who posits the following constraints:

(11) Faithfulness constraints pertaining to lexical stress

MAX(LA): A lexical accent in the input has a correspondent in the output

DEP(Root): A lexical accent in the output has a correspondent in the input

Stress placement for a word for which lexical stress marking has to be assumed, namely *ʒe'nōh* ‘enough’ is illustrated in Tableau 7 below. In candidate (a), the lexical stress on the

⁵² That grammatically assigned stress can co-exist with occasional lexical marking for stress is illustrated by Polish, where despite the existence of highly regular penultimate word stress (cf. such surprising, from a non-Polish perspective, forms as *in'ternet*), lexical exceptions are allowed, such as *'cztery* ‘four hundred’ (cited by Sussex and Cubberley 2006: 16) or, increasingly, *'prezydent* ‘president’, both with initial (or: antepenultimate) stress.

second syllable of the input form has no correspondent in the output, hence a violation of MAX(LA) is incurred. Additionally, in the same candidate, the lexical stress on the initial syllable of the output form has no correspondent in the input, hence a violation of DEP(LA) is incurred. This illustrates the fact that a case of ‘shifting’ of lexical stress of a given syllable of the input onto a different syllable in the output would amount to a simultaneous violation of MAX(LA) and DEP(LA). It is clear that MAX(LA), (DEP(LA)), or both must have dominated TROCHAIC. In other words, as manifested in a considerable number of words, it was more important to keep the stress marking in the surface form in line with that of the input form, than it was to have trochaic feet.⁵³

<i>ʒeˈnōh</i>	MAX(LA)	DEP(LA)	TROCHAIC
(a) óσ	*!	*!	
(b) σό			*

Tableau 7: Stress placement in lexically marked forms

4.3.4 Summary

As a result of the promotions and demotions, the duration of stressed vowels was often adjusted and became variable. Consequently, duration was no longer a reliable cue to the phonemic identity of the vowels in question. As is suggested in the next section, the job of keeping the long and short vowels apart was taken over by qualitative differences.

4.4 EARLY MODERN ENGLISH

The Early Middle English period witnessed the events which form the starting point for the discussion of vowel chain shifting in this thesis, namely the Great Vowel Shift. The description of these historical events, together with the discussion of some of the many contentious points related to them is presented in Section 3.1.1 above. Here, an argument is

⁵³ Forms such as OE *ærende* ‘errand’ attest to this happening in OE already. However, with the growing number of items where morphological transparency was lost, such cases would have become much more widespread in ME.

presented that the GVS, and especially its early stages, are closely linked with the weakening of the role of duration for the maintenance of phonemic vowel contrasts.

4.4.1 The replacement of length with tenseness

From OE through Middle English, phonetic duration was, primarily, an expression of the phonological feature of length. In light of the view of OT adopted here, under which constraints are not innate but constructed by learners, this means that the distinctive role of length would lead language learners, when exposed to ambient language, to construct constraints which make reference to length. One of them is the IDENTIO(LENGTH) constraint. Another is PC(length). Whereas some degree of predisposition to set up constraints probably has to be assumed, and maybe even a more specified predisposition to set up faithfulness or PC constraints, the precise instantiations of those can be stipulated to emerge with exposure to ambient language.

Building on the PC family of constraints introduced in 2.3.2.2 above, I suggest that PRESERVECONTRAST(P) ((PC(P)) constraint is constructed by learners and that it is a force against mergers in the process of historical change. PC(P), militates against collapsing phonemic categories maintained by a phonological feature for a given segmental feature. This constraint is set up for any feature for which there is sufficient support throughout the phonemic inventory. Given the recurring cases of quantity based contrasts being re-coded as quality-based contrasts instead of being lost completely (in the developments Latin > Common Romance, Sanskrit > Hindi, Old Icelandic > Modern Icelandic, early Proto-Slavic > late Proto-Slavic mentioned in Section 3.4 above), it seems that there is a pressure against wholesale mergers of phonemes that used to be kept apart by a feature that is now disappearing from the system. Instead, contrasts which are maintained by a vanishing feature, are kept apart by other means, be it by means of the features already present in the system (e.g. in the developments referred to above, where quality-based oppositions were re-coded as quality-based, with possible introduction of new vowel-height degrees), or by the introduction of a new feature, replacing the vanishing one (e.g. in the account proposed for English here, where length-based distinctions are argued to have been re-coded as [±tense]-based). Naturally, the presence of this constraint is not meant to imply that all contrasts which are endangered by a disappearing feature will be maintained. Like all constraints, it is

violable, and is merely an embodiment of the observation that the seemingly most obvious consequence of the disappearance of a feature, i.e. the collapse of vowel pairs previously kept apart by this feature, is militated against.

(12) PRESERVECONTRAST(P) (PC(P): Do not merge a contrast maintained by feature P.

There is an important difference between a simple identity constraint referring to a given feature and a PC constraint referring to a contrast based on that same feature. With regard to length, for instance, while IDENTIO(LENGTH) assigns a violation mark for each output segment which differs in its specification for length between the output and the input, PC(LENGTH) assigns a violation mark for each pair of phonemes which are distinct in the input but non-distinct in the output. If that pair of phonemes is kept distinct in the output by a different feature or a set of features than it is in the input, no violation marks will be assigned.

Naturally, an important question is, under what conditions is it legitimate to postulate that a PC constraint for a given feature is constructed by learners. It seems safe to assume, though, that a feature neatly cutting right through the entire vocalic system of a language (which used to be the case with Length in OE) is as good a candidate as possible. There is a sense in which the issue of whether there is sufficient evidence in the ambient language to posit a relevant PC constraint for that feature is reminiscent of the discussion of functional load (Martinet 1952, for an investigation of the explanatory value of predictions made by functional load for language change, cf. King 1967, for further criticism see Lass 1980 [2009]: 91). The feature-based view on the preservation of contrast as embodied in the PC(P) has a different focus than that of functional load, since, in its diachronic manifestation⁵⁴, it pertains to cases where a feature is being lost from the phonology of a language, which then results in lexical mergers. As is well known, mergers do occur. Some were even implicated in the events immediately following the GVS, namely the merger of MEET and MEAT words, i.e. the raising of /e:/ in *meat* to join the already raised /i:/ in *meet* in the eighteenth century (Lass 1999: 96). Still, in cases like this one, though a PC(P) constraint (here: PC(HEIGHT)) is violated by this relevant pair of phonemes, it is not violated by all vowel

⁵⁴ PC constraints were devised by Lubowicz (2003) to handle contrast preservation in synchronic input/output mappings. In the following analysis, they are applied to model diachronic developments.

pairs maintained by height. There are two predictions made by the view of PC(P) constraints presented here for historical developments. First, it can be informally stated that for PC(P) a merger of one phonemic opposition is better than a merger of more phonemic oppositions, since multiple merger equals multiple violations. Second, since for a PC(P) constraint to be constructed in the first place enough evidence in the ambient language must be provided, the prediction is that mergers should be more prevalent among phonemes kept apart by features with limited presence in the phonological system in question. The predictions made by following the functional load perspective are slightly different. There, the likelihood of a particular merger is assessed against the likelihood of another merger happening, and the assessment is conducted on the basis of the lexical contrasts that the respective oppositions help maintain. Because of those differences in the kinds of predictions made by the importance of PC(P) constraints and by functional load, King's (1967) study does not preclude the possibility that PC(P) constraints are operative in the avoidance of mergers. Still, it does provide a suggestion about the methodology that could be applied to test it.

Around the time when forms ambiguous with regard to their duration were arising, the presence of the PC(LENGTH) constraint ensured the persistence of the contrasts. They were not collapsed, however unreliable they were to maintain. It is assumed that learners were in effect compelled not to posit input forms with the lexical sets merged, because they had sufficient evidence to set up two separate underlying representations. It is further assumed that they did not merge these contrasts in production because, on the basis of the same evidence, they constructed relevant PC constraints. As speakers, they kept the relevant lexical items apart, even though they did not necessarily distinguish them by length anymore. Learners exposed to input which kept the contrast, though not by length, were in turn compelled to set up inputs containing the contrast, and a concomitant PC constraint, this time one for tenseness, rather than length.

This argument can be seen as functional, in the sense that, from a speaker-based perspective, it would seem to suggest that speakers maximize the contrast for it to better serve their communicative needs, thus making their language more functional. A counter-argument levied by Lass (1980 [2009]) against such an appeal to functionalism is that it leads to the untenable conclusion that successive generations of speakers display increasing need for functionality, and here is how he argues. In a speaker-based, functional account it is

assumed that a functional change happens because speakers want to improve their language. Since one generation (Generation₁) is able to content itself with the state of language as it is, and a Generation₂ introduces a change which makes their language more functional, it must be concluded that the reason why the next generation (Generation₂) comes to implement the change is that the need for functionality is higher among speakers of Generation₂ than it was for Generation₁ speakers. For instances of language change which span several generations of speakers, this need for functionality would have to grow successively with each generation. Thus, a functional, speaker-based account of a change like this one, that is of maximization of contrast, would actually replace the necessity to explain the spread of a new kind of linguistic system (a system with greater contrasts than the minimal contrasts of the initial stage) through the population of its speakers, with the necessity to explain the spread of the degree to which functionality is valued by the population of speakers (i.e. the spread of the need to have maximized phonemic contrasts instead of having minimal phonemic contrasts). Thus, the change in the language is not explained, but the task becomes to explain the change in preferences of its speakers. More importantly, such a progressive increase in the need for functionality in language seems hardly tenable, and so nothing has been gained by bringing in functionality in the first place.

From the evolutionary perspective, however, the fact that an unreliable, or ‘weak’ contrast can exist and function in one generation, is fully compatible with the fact that it becomes maximized over successive generations. In biological evolution, individual organisms of Generation₁ are also viable, just like a language with an unreliable or ‘weak’ contrast is ‘functional’. Still, biological features can change, or be optimized, even when no environmental changes are involved. Likewise, the realization of an already ‘functional’ contrast, that is of a contrast which is capable of maintaining the distinctness of lexical sets, can maximize over time

4.4.2 The setting off of the GVS

Before the changes that took place in EModE are presented, the pre-GVS, Middle English manifestation of length-based contrast has to be introduced. This initial stage for vowel pairs /e:/, e/ and /o:/, o/, which are assumed to be the first in the long line of vocalic changes starting

in the EModE period, is presented in Tableau 8 below. For speakers with such grammars, the contrast is specified in the input as one of length, that is vowels are specified as either having to attach to one or to two nodes on the tier. Additionally, the usual featural faithfulness constraints apply. The MINDIST constraint, requiring that contrasts are maintained with sufficient robustness, is satisfied by all candidates at this stage, but is introduced here because of its importance in the following stages.

	PC(LENGTH)	MINDIST	IDENTIO(HEIGHT)
be:t, bet ↗/be:t, bet/			
/bit, bet/			*!*
/be:t, be:t/	*!		
/bet, bet/	*!		
bo:t, pot ↗/bo:t, pot/			
/but, pɒt/			*!*
/bot, pot/	*!		
/bo:t, po:t/	*!		

Tableau 8: Stage 1 (pre-EModE, contrasts coded as LENGTH based)

As outlined in 4.3.2 above, due to the metrical adjustments on the post-lexical level, this underlying length contrast was not consistently realized as a durational contrast in surface forms. Thus, due to the fact that phonetic duration was beginning to be the expression of prosody, it did not function very effectively as a cue to segmental contrasts. Even though the contrasts were maintained, they were minimal. However, due to the presence of PC(LENGTH), the output forms merging the contrasts were ruled out. Instead, the contrasts were maintained by the accumulated qualitative differences. These gained the upper hand over the length-based realizations which came to violate MINDIST, since even though the contrast was maintained, it was too small. Due to this violation of MINDIST, the output forms which manifested qualitative differences, previously sub-optimal due to violating IDENTIO(HEIGHT), came to be optimal.

	PC(LENGTH)	MINDIST	IDENTIO(HEIGHT)
be:t, bet		*!	
ɪ̥bet, bet			*
be:t, be:t	*!		
bet, bet	*!		
bo:t, pot		*!	
ɪ̥but, pɒt			*
bot, pot	*!		
bo:t, po:t	*!		

Tableau 9: Stage 2: (transitional, contrasts coded as LENGTH based, realized by means of quality)

Following generations of learners began maximize the phonetically motivated qualitative differences between the members of the pairs. Those exemplars were more successful which were successively more distinct from each other, which fact found its way into the phonetic component.

As these differences grew, a qualitative shift happened, in which in some speakers the linguistic input led to the TENSE constraints being set up instead of the LENGTH constraints as it used to be the case in more conservative speakers. Given the rather scarce evidence for a duration-based contrast and an ever more robust evidence of a quality-based contrast, the language systems of the innovative speakers had the following characteristics. First, they posited tenseness-based contrasts in the lexicon. Second, the relevant PC constraint came to be PC(TENSE), as opposed to PC(LENGTH) of the resident grammars. Those two differences must be thought of arising at the same time, since the same sorts of characteristics of the ambient language which are adduced to set up constraints, here, the prevalence of constraints based on tenseness, are also at play when lexicon entries are established. This post-reorganization stage is represented in Tableau 10 below.

	PC(TENSE)	MINDIST	IDENTIO(HEIGHT)
/bit, bɛt/ be:t/bet		*!	*
↗↘bit/bet			
be:t/be:t/	*!		*
bet/bet	*!		*
/but, pɒt/ bo:t/pot		*!	*
↗↘but/pot			
bot/pot	*!		*
bo:t/port	*!		*

Tableau 10: Stage 3 (EModE, contrasts re-coded as TENSE based)

Under this account, no re-ranking of constraints was involved in the change. Rather, the content of one constraint, the PC(P) constraint, changed from one generation to the next; PC(LENGTH) was replaced with PC(TENSE).⁵⁵ At the same time, the featural specification of lexical entries changed; again, length was replaced with tenseness. Since the motivation for both, the content of the PC constraint and the features posited for the input lies in the nature of the ambient language, it is not too far-fetched to assume that such a re-structuring in the lexicon and in the content of a constraint can occur simultaneously. This instance of linguistic change cannot be presented within a tableau, since an instance of language change is not a case of an input/output mapping, with the resident form being the ‘input’ and the new form being the ‘output’. All that Tableau 10 can show, then, is that, assuming there was enough evidence in the ambient language to posit tenseness-based contrasts and a relevant PC constraint, neither the old contrast nor a merger would survive the heat of the contest. The representation of the phonologized accumulated effects of strengthening under stress remains elusive.

The issue of the ‘residual’ durational differences needs some commenting on. Admittedly, the PDE tense vowels are, when in the same context, longer than PDE lax vowels, so *beat* is still longer than *bit* (even though *beat* is shorter than *bid*). The durational

⁵⁵ Naturally, if one decided to stand by the innateness of constraints, one could argue that PC(TENSE) initially occupied such a low position in the ranking that it was non-operative, that it was then promoted to the position occupied up to that point by PC(LENGTH), which, in turn, was demoted very low indeed.

characteristics that result in this difference must therefore be assumed to be a part of the respective phonemes, even though the primary feature distinguishing them is, phonologically, tenseness and, phonetically, quality. Thus, the situation is the obverse of what used to be the case before EModE. Just as for pre-EModE, it has to be assumed that qualitative differences accompanied quantitative differences, and so non-distinctive characteristics of sounds were part of their identity, so it is the case now that durational differences accompany qualitative differences.

This initial development, that is the qualitative differentiation of ME /e:/ and /e/ into EModE /i:/ and /ɛ/, as well as of ME /o:/ and /o/ into EModE /u:/ and /ɔ/, set off the GVS. The raising of ME /e:/ and /o:/ encountered ME /i:/ and /u:/ on their paths, and so the latter diphthongized. The raising of mid-high vowels and diphthongization of high vowels is presented in Tableau 11 below.

	PC(HEIGHT)	IDENTIO(HEIGHT)
beet, bite /bit, beɪt/ ☞ (a) /bit, beɪt/		
(a) /bit, bɪt/	*!	*
(b) /bi:t, bi:t/	*!	*
(c) /be:t, bi:t/		*

Tableau 11: The raising of mid-high vowels and diphthongization of high vowels

The absence of one candidate from Tableau 11 warrants some explanation. Once ME long /e:/ became EModE tense /i/, one could be tempted to posit that it was distinct from the ME /i:/, in that the latter was still specified as long, rather than tense. This set of contrasts is presented below as (13).

- (13) beet, bite
 /bit, bɪt/
 [bɪt, bi:t]

Such a set of underlying forms cannot be constructed by learners, though. Since phonological tenseness still entailed phonetic duration, the two vowels would have been indistinguishable

on the surface. Consequently, there would be no evidence to construct such an opaque complex of distinct vowels, two distinctive features, which would yield a collapsed contrast.

Again, the actual change can by no means be seen happening within a tableau, as the instance of linguistic change is a difference between two competences and not a mapping of the older form and a newer form within a single competence. The fact the ME /i:/ developed an inglide and thus changed into /ei/ cannot be represented as input/output mapping of a single competence. The development of an inglide is a result of the accumulated effects of strengthening under stress. This effect must have become stabilized, so that the ambient language in which the language acquisition of the following generation took place provided evidence for positing a diphthongal underlying form.

The abundance of data in the ambient language resulting in the construction of the PC(HEIGHT) constraint cannot be doubted, since height contrasts are one of the most basic in vowels. Its existence, in conjunction with the presence of a height difference between ME /e:/ and ME /i:/, namely the presence of a non-high element in the inglide of ME /i:/, ensured that the contrast was not merged.

4.5 PRESENT DAY ENGLISH

The vocalic system of PDE, a result of the developments described above, is argued to contain within it the ‘shiftiness’, which is characteristic of many of the Present-Day varieties of English. Even though the main analysis is focused on one lineage of English, one whose current instantiation is RP, a brief comparison to relevant developments in two other varieties (New Zealand English and Australian English) is also included at the end of this section.

The developments in the phonological system of English presented so far have resulted in a vocalic system which, in light of the study reported on in Section 3.2, seems to be more prone to vowel chain shifts now than it used to be before the EModE times. This could be so for the following reasons. First, there are more distinctive vowel qualities nowadays than there used to be; there were nine vowel qualities in the tenth century whereas there were nineteen in the twentieth century. The results of the study reported on in 4.5.1 below add to

the argument that tense/lax vowel pairs, such as /i:/ vs. /ɪ/, are differentiated primarily by means of quality rather than quantity, which justifies postulating distinctive qualities for members of such pairs. The data gathered for that study was further analyzed with respect to the co-variation of duration and quality. This analysis, presented in 4.5.2, suggests that qualitative variation is not predictable on the basis of durational variation. The reasoning behind it is that since duration is under the control of so many phonological and phonetic factors, it can no longer be nicely correlated with quality. As a result, the reliance on minute details of quality is heavy. In other words, qualitative variation, not being easily predictable from quantitative variation, is salient, so that very small qualitative differences are paid attention to. At the same time, stress placement is relatively stable (though not as stable as in OE, since in PDE there are derivational affixes that shift stress, i.e. that eliminate the primary stress of one form altogether, e.g. 'symbol – sym'bolic), which continues to allow for the entrenchment of the segmental effects of stress. Lengthenings and shortenings are not available to be exploited for social purposes, since duration is entangled in the expression of a number of factors other than inherent vowel length. A development observed in the twentieth century, which is likely to make it impossible for vowel length to become a primary vowel feature distinguishing numerous vowel pairs in English, is that the duration of a vowel is under strong influence of the sonority of the following consonant, so much so that the vocalic length contrast may be suspended, or even reversed. Since the need to mark social affiliation through the use of particular language variety, which since French lost in status in England could no longer be achieved by the simple choice of French vs. English (Perkins 1977; Leith 2002; Lerer 2007) has arguably remained undiminished, the conditions for continued vowel shifting are still in place.

4.5.1 Study 2a: Quality is the main expression of vocalic contrasts in PDE

4.5.1.1 The changing roles of vowel duration and vowel quality for the expression of phonological contrasts

There are languages of the world in which the physical duration of vowels is the main correlate of the phonological feature length. In a language like Czech (Dankovičová 1999), for instance, the phonological contrast between several vowel pairs is maintained primarily

by means of duration, with any possible differences in quality being of secondary importance. English, however, is not one of those languages. Although it used to be one, possibly up to the Late Middle English times, now vowel contrasts are maintained primarily by means of vowel quality, with any possible differences in duration being of secondary importance.

Admittedly, there are still remnants of the old system in which the weight of the contrast rested with duration. For instance, all other things being equal, /i:/ is still longer than /ɪ/, that is a vowel pair which used to differ only in duration, and hence constituted a contrast based on length, still displays a difference in duration. However, the stipulation ‘all other things being equal’ is crucial here, since when other things are not equal, namely, when the vowels in question are followed by consonants differing in voicing, duration not only ceases to mark the contrast, it even points in the other direction. And so the historically short vowel /ɪ/ is phonetically longer than the historically long vowel /i:/, when the former is followed by a voiced consonant, and the latter by a voiceless one (Gimson & Cruttenden 2008). A further fact undermining the role of duration for maintaining phonological contrasts in English is that, even when other things are equal, a historically long vowel is not necessarily longer than a historically short vowel, with /æ/ being actually longer than /i:/ or /u:/ (House & Fairbanks 1953: 111). Perceptual acoustic research (Clark & Hillenbrand 2003: 10ff.; Bogacka 2003) corroborates the primacy of spectral cues over temporal cues for the identification of vowels by listeners. That speakers rely mostly on quality when identifying vowels has been found for high vowels, potentially the last strongholds of the temporal distinction. That is, /i:/ and /ɪ/, as well as /u:/ and /ʊ/, would be the best candidates for the realization of a length contrast, since members of these pairs do not differ dramatically in quality and they do differ, albeit only in the same phonetic context, in terms of duration. The fact that even for their identification spectral cues are more important than temporal cues, makes an even stronger case for the importance of spectral cues relative to temporal cues for the phonological identity of English vowels.

This is how the loss of phonemic vowel length could have contributed to the onset of vowel shifting.

Until the Early Modern English period two classes of vowels were distinguished, phonologically by length and phonetically by duration. As argued already in 3.4, these quantitative contrasts may well have been accompanied by qualitative differences. However, as long as the relationship between the durational variation, and qualitative variation was straightforward, any qualitative differences that might have obtained between the members of the pairs could be attributed to the primary quantitative difference. As a result of the developments between the fifteenth and the seventeenth centuries (laxing of short vowels, raising and diphthongization of long vowels), contrasts maintained mainly by means of duration were now expressed mainly by means of quality. The differentiation of long and short counterparts into contrasts increasingly differing in quality involved the changes underwent by the long vowels known as the GVS. The newly established pairs have not come to be distinguished by length, and so the pre-GVS balance has never been restored.

The differentiation of short and long vowels into qualitatively different phonemes had two important consequences. First, from the perspective of the vocalic system itself, the number of qualitatively distinct vowel phonemes has substantially increased, and this in itself might have contributed to the likelihood of shifts. More distinctive vowel qualities means more potential candidates for a shift to start. Also, an increased number of distinct vowel qualities results in an increased number of potential ambiguities, and so individual vowels can move to find their respective niches. Second, as the previously qualitatively identical vowels were drifting apart, a movement was set in motion that was then to continue, because the conditions for it still hold, in PDE more so than ever with the additional influence of coda voicing on vowel duration. The very fact that new vowel qualities were now becoming distinct meant that the acuity of the listeners to vowel quality must have become more fine tuned. In other words, the increase in importance of subtle, previously irrelevant qualitative differences raised the attentiveness of the speaker/hearers to small qualitative differences. Through OE and ME, when durational differences were the primary expression of phonological contrasts, the possible qualitative differences could be predicted from them. So if a realization of a vowel were a little less peripheral than its other realizations, it was more often than not the case that it was also shorter, and this allowed its reconstruction as an inherently peripheral vowel. For instance, up until the fifteenth century, when vowel length was phonemic, the subtle qualitative differences between /i:/ and /i/, /e:/ and /e/, /a:/ and /a/,

/o:/ and /o/, /u:/ and /u/ could have been filtered out, since they could, most of the time, be put down to contextual factors. That is, for example, short /e/ and /o/ were more central than long /e:/ and /o:/ because they were shorter than them. Conversely, the long /e:/ and /o:/ were more peripheral than /e/ and /o/ because they were longer. Even if the correlation between duration and peripherality did not hold for each and every production, it is likely that it held most of the time, and so a generalization could be derived by listeners. Such contextually induced lengthenings and shortenings could have been filtered out by listeners just as the nasality of a vowel followed by a nasal stop is filtered out by English listeners nowadays, along the lines of the ‘compensatory correction’ referred to by Ohala (1981: 187). As suggested above, this correction became increasingly difficult, however, as phonetic duration became increasingly interfered with for rhythmic purposes. In avoidance of stress clashes and lapses, English vowels were promoted/demoted or stretched/compressed. This would mostly happen to unstressed vowels. However, when a number of them were lost altogether and the number of monosyllables rose, stressed vowels were increasingly subjected to these strategies. As a result, a promoted short vowel could have similar duration to a demoted long vowel, or a stretched short vowel could have a similar duration to a compressed long vowel. Consequently, as the stretching and compression of lexically stressed vowels due to rhythmic constraints on the post-lexical level became more widespread, and the phonemic /e:/ and /o:/ were still more peripheral than /e/ and /o/, even though they were very often phonetically of no longer duration anymore, listeners were left wondering as to the reason behind these qualitative differences. As soon as the variation in quality could not be predicted from the variation in duration, i.e. as soon as a vowel is less peripheral than its other realizations, but is not shorter than those other realizations, inherent qualitative differences have to be postulated for that vowel. The qualitative differences were now interpreted as intended, just as, say, contextually nasalized vowels were interpreted as intentionally nasal at the time of the loss of the following nasal stops in French at the turn of the second millennium. Once such qualitative differences were interpreted as intended, they could be exploited for the purpose of expressing social group affiliation.

For the duration-induced qualitative differences to be picked up on by listeners, they must have become part of the phonemes themselves, or part of language-specific phonetics. If a qualitative difference had been produced on-line for each production of a vowel, then the qualitative differences would have been absent in each production with other than the underlying length. The subtle qualitative differences must be allowed to become entrenched in the representations of the relevant phonemes over long stretches of time, for them to remain available once the conditioning factor is gone. It is hypothesized that the qualitative differences had become stabilized by the Early Modern English period. Since the conditions which caused qualitative difference to not be predictable from quantitative still hold, the rearranged vowel system of post-GVS English has not fallen into a symmetrical length-based pattern ever since. What stands in its way is that vowel duration is still contingent on a number of factors besides its inherent value, now, at least since mid twentieth century (House & Fairbanks 1953) also on the voicing of the coda consonant. Although the rearranged post-GVS vowel system does display vowel pairs of similar quality, namely the tense/lax pairs of /i:/ - /ɪ/ and /u:/ - /ʊ/ , their qualities are definitely very much kept apart.

4.5.1.2 The study

Eight subjects, four female and four male, were recorded reading four different materials. Each test word contained one of four front vowels, namely /i: ɪ ɛ æ/ in stressed position. Each vowel was placed in three different words: a one-syllable word with a voiced coda, a one syllable word with a voiceless coda and a two syllable word.

First, subjects read a list of carrier sentences with test words embedded in stressed, word final position. Then, they read the same list but were asked to speed up their performance and read at the highest possible comfortable pace. Next, they read a short narrative passage containing four of the test words (here, only one syllable words with a voiced coda were used). Finally, they read a list containing isolated test words, where each word appeared twice. The entire material that was administered to the subjects can be seen in Appendix 3. This amounts to 52 measurements per speaker – four vowels multiplied by three words multiplied by four repetitions (once in each reading of the list of sentences and twice in the

reading of the list of isolated words) plus four (words in the narrative passage). With eight speakers, a total of 416 tokens of test words were present in the material.

The responses were recorded using a Roland portable recorder and saved as .wav files. They were then analyzed using the Praat (Boersma & Weenink 2009) software. The following measurements were taken by hand: duration and the frequencies of the first two formants (F1 and F2) in Hz for the mid-points of the steady states of the vowels. To enable the measurement of duration, the widely accepted criteria for segmentation of Peterson and Lehiste (1960) were followed.

4.5.1.3 Analysis

To investigate the roles of quality and duration for the expression of the identity of vowel phonemes two analyses were conducted. The first one assumes a simplified bottom-up nature of speech perception, and the second one takes into account the parallel processing of units of different levels in speech perception.

Under the bottom-up view, the units at the lowest level of linguistic organization are interpreted first, and then the results of this interpretation are fed into the next level. Consequently, for the discrimination of phonemes no higher-order units can be invoked. Hence, for the durational differences to be primary, with the qualitative differences being predictable from them, the following would have to hold.

- (14) The relation of the quantitative difference between /i:/ in *beat* and /ɪ/ in *bit* to the qualitative difference between /i:/ in *beat* and /ɪ/ in *bit* should be the same as the relation of the quantitative difference between /i:/ in *beat* and /i:/ in *beating* to the qualitative differences between /i:/ in *beat* and /i:/ in *beating*.

In other words, since under the serial view listeners cannot use the information about foot structure before they have processed the individual phonemes, they cannot treat the stressed vowel phoneme of *beating* any differently *than* the stressed vowels of *beat*, *bit* since they have no reason to suspect the influence of foot length on the duration of the former. To interpret the phoneme, then, they rely on phonetic duration only, and the phoneme with the shorter duration will differ in quality, too, regardless of whether the phoneme is shorter due to inherent durational contrast (*beat* vs. *bit*) or due to being part of a two-syllable vs. one-

syllable foot (*beat* vs. *beating*). The extent of the concomitant qualitative difference is expected to be the same. Since listeners cannot be expected to make use of the information about foot structure, speakers could be expected not to make them.

To test this prediction, mean duration, F1 and F2 of all tokens of *beat*, *bit* and *beating* were pooled. These can be seen in Table 7 below.

	Duration (in ms)	F1 (in Hz)	F2 (in Hz)
<i>beat</i>	129	365	2499
<i>bit</i>	96	486	2160
<i>beating</i>	98	365	2498

Table 7: Mean duration, F1 and F2 for *beat*, *bit* and *beating*

As can be seen in the table, the prediction following from assuming the primacy of durational differences under the bottom-up view of speech perception is not fulfilled. The prediction in (14) cannot be met since, although the durational difference between *beat* and *bit* on the one hand and *beat* and *beating* on the other hand is comparable, there seems to be no qualitative difference between *beat* and *beating*, while there is one between *beat* and *bit*. A statistical analysis comparing effect sizes, which takes variance into account, confirms it. For the pair *beat/bit*, both the difference between mean durations ($t=5.69$, $df=31$, $p=3\times 10^{-6}$) and the difference between mean F1 values ($t=-8.58$, $df=31$, $p=1\times 10^{-6}$) are statistically significant with large effect sizes ($d=1.42$ and $d=-2.14$ respectively).

For the pair *beat/beating* however, while the difference between mean durations is statistically significant with large effect size ($t=6.17$, $df=31$, $p=8\times 10^{-4}$, $d=1.54$), the difference between mean F1 values do not reach statistical significance ($t=0.003$, $df=31$, $p=0.99$), and the effect size was estimated at $d=0.000675$, which is smaller than small according to Cohen's (1992) classification. A post hoc power analysis reveals that, assuming that this effect size reflects the true effect size of the population, a total of 565468 pairs would have to be measured to find a significant difference ($\alpha=0.05$), assuming there is one. Not only would conducting such an experiment not be practical, but it is also doubtful that such a small effect size would be of any practical relevance. It could therefore be said that the relation in (14) does not hold. The relation of the quantitative difference between /i:/ in *beat* and /ɪ/ in *bit* to the qualitative difference between /i:/ in *beat* and /ɪ/ in *bit* is that of a large

difference to a large difference, whereas the relation of the quantitative difference between /i:/ in *beat* and /i:/ in *beating* to the qualitative differences between /i:/ in *beat* and /i:/ in *beating* is that of a large difference to, most likely, no difference, or a small difference.

This analysis may be questioned on the grounds that it relies on a linear view of phonological processing, in that it disregards the context of other constituents next to which these vowels occur. Indeed, a large body of work in formal phonology, brought together under the umbrella term ‘nonlinear phonology’ has shown that the way in which phonological representations are translated into speech are very much dependent on the actual sequence of constituents within which they are embedded (for an overview see e.g. McCarthy 1982, Pulleyblank 1989). Aside from formal work, the nonlinear view of phonology has also found support from perceptual studies, where it has been shown that listeners in fact process the incoming speech signal on different levels of structure in parallel rather than serially (Hawkins 2003, 2010). Hence, constituents which are higher up in the hierarchy, say syllables or feet, can and do influence the processing of constituents which are lower in the hierarchy, say individual segments. Thus, the kind of foot in which a given phoneme resides is not irrelevant when a listener decides upon its identity.

Bearing that in mind, it is still worth investigating how phonemic contrasts are actually maintained in production. Mean values for /i:/ and /ɪ/ are presented in Table 8 below.

	Duration (in ms)	F1 (in Hz)	F2 (in Hz)
bead	239	349	2547
beat	129	365	2499
beating	98	365	2498
bid	147	452	2194
bit	96	486	2160
bidding	72	433	2126

Table 8: Mean values for duration, F1 and F2 for /i:/ and /ɪ/

It can be seen that F1 and F2 means for *bead*, *beat*, *beating* are all higher than the means for *bid*, *bit*, *bidding*. On the other hand, durational means for the two phonemes are not kept apart, and the words can be seen to follow the following hierarchy: *bead*, *bid*, *beat*, *bit*=*beating*, *bidding*. Considering the consistency in maintaining the contrast by means of quality, compared to a lack of such a consistency with regard to duration, it makes sense to

assume that the qualitative differences are primary. It is certainly more parsimonious to assume that the quality of a vowel is attended to as the primary cue, since it suffices to determine each sound's identity, than to postulate that duration is relied upon, in combination with the information about foot structure.

These results point to the primacy of the qualitative characteristics of vowels in the expression of phonemic oppositions. These are additionally enhanced by duration, but in view of these results the role of quality is greater than that of duration. All this seems to corroborate the phonological account of Giegerich (1992), seeing [\pm tense] as the primary feature distinguishing between two classes of vowels in English, with [\pm length] predictable from it.

4.5.2 Study 2b: Qualitative variation is not predictable from durational variation in PDE

The hypothesis entertained in this thesis is that the increased likelihood of the English vowel system to undergo vowel chain shifts is, in part, a result of the loss of duration as a feature distinguishing two sets of vowel phonemes from each other. It is further hypothesized that the loss of length as a distinctive feature in vowels was brought about by the fact that qualitative variation ceased to be predictable from quantitative variation. The analysis of data presented in the following sections seeks to establish whether a lack of a tight-knit relation between qualitative and quantitative variation is still a characteristic of English nowadays. The results seem to indeed agree with this assumption. This unpredictability of qualitative variation from quantitative variation is argued to contribute to the diachronic instability of the vocalic system, as manifested in the recurring vowel shifts in PDE.

This lack of reliability of duration to signal phonemic oppositions in vowels, which is argued above to have contributed to qualitative splits between pairs of vowels previously distinguished by length, has remained a feature of English to the present day. Just as initially in EModE long and short phonemes of similar quality started drifting apart because their contextually-induced qualitative differences were no longer predictable from their durational characteristics, so it is the case today that the qualitative variation within individual phonemes cannot be filtered out, and so becomes entrenched.

In the following, a study into the relation between variation in duration and variation in quality is reported on, following an overview of extant research on variability of vowel duration and vowel quality.

4.5.2.1 The variability in vowel duration

The phonetic duration of a vowel is influenced by a range of factors which can be broadly divided into four groups: the intrinsic duration of the vowel itself, its phonetic context, suprasegmental features and syntactic features.

In terms of the nature of the vowel itself, it has been observed that English vowels possess intrinsic duration, and they can be divided into intrinsically long /i:, eɪ, æ, ɑ:, ɔ:, ou, u:, ɜ:, aʊ, ɑ:r, ɔ:r/ and intrinsically short /ɪ, e, ə, ʊ/ (Peterson & Lehiste 1960: 703). One noteworthy aspect here is that the phonetically long and phonetically short classes do not completely coincide with the classes of phonologically long and short vowels of ME, as /æ/ is durationally long. At the same time, it patterns phonologically with other vowels which used to form the class of short vowels in ME, which undermines the case for length as a feature characterizing English vowels.

It can also be noted that more open vowels were found to be longer than less open vowels (Delattre 1962: 1141). The motivation for this, however, is not entirely clear. There are physiologically oriented explanations which invoke the greater time required for the articulation of low vowels as a reason for open vowels being longer. However, such accounts fail to take into account that low vowels do not display greater onglide to steady state duration ratios compared to high vowels (Lisker 1974: 236). In other words, if open vowels simply took longer to articulate, it would be expected that they have a longer onglide stage than close vowels, that is that their articulatory goals are reached later. This, however, is not the case.

Furthermore, the correlation between vowel height and duration holds only within the class of intrinsically long vowels. In line with it, /i:/ and /u:/ were found to be shorter than /eɪ/ and /ou/, which were in turn shorter than /æ, ɔ: and ɑ:/. When it comes to short vowels,

however, no systematic link between vowel height and duration was found (Rositzke 1939: 103). Further, the two findings are curious with regard to the predictions that would derive from them for historical developments. These stem from comparisons of vowels across the classes of intrinsically long and short. Here, the correlation between duration and vowel height is reversed, with the intrinsically long vowels in each pair, e.g. /i:/ compared to /ɪ/, being systematically both less open and longer (House 1961: 1177). These two English vowels used to have a very similar quality, presumably /i/. Now, if shorter vowels are less open, then the shorter of the pair should be less, not more open, than the long counterpart, and the entrenchment of this difference should not have resulted in a lowering, but in raising. However, ‘intrinsically’, or phonologically short counterparts in each such pair lowered. The example of /i:/ and /ɪ/ is paralleled by the developments of /u:/ and /ʊ/, /e:/ and /ɛ/, and /o:/ and /ɔ/ in EModE.⁵⁶ It seems, then, that historical developments of phonologically short vowels (i.e. laxing or lowering) are subject to different processes than those responsible for phonetic durational differences between the vowels within the class of phonologically long vowels.

All in all, even though duration, as argued above and supported by the results of the study reported on in Section 4.5.1, is not crucial for determining vowel identity in English, vowels have some intrinsic durational characteristics, the exact nature of which remains to be explained. Whereas the hierarchy within the class of tense vowels may have its origins in articulatory facts, it must have already become part of language-specific phonetics. This would suggest that low vowels tend to become longer over time. Further, the reversal of the correlation for tense/lax pairs, paradoxically suggests that short vowels become lower.

With regard to the phonetic context, the following factors affecting vowel duration can be named: the voicing, manner of articulation and, to a certain extent, the place of articulation of the following consonant. Studies probing the influence of adjacent consonants on the duration of vowels agree unequivocally that English vowels are shorter before unvoiced than before voiced consonants (Rositzke 1939; Lehmann & Heffner R-M. S. 1943; House & Fairbanks 1953; Peterson & Lehiste 1960; House 1961; Chen 1970). It is the most widely

⁵⁶ It is also paralleled by the developments in other languages discussed in Appendix 1 below.

attested environmental effect, and it also, arguably, the one having the greatest impact on the phonology of English, as the extent of its effect is often cited as an argument for the primacy of quality over quantity in the specification of English vowels (e.g. Giegerich 1992). The study reported on here investigated only the influence of voicing of all the contextual factors.⁵⁷

As far as prosodic factors are concerned, the stress of the syllable whose nucleus a vowel happens to be, and the length of the foot it occurs in have an impact on duration. Duration has been found to be one of the correlates of stress in English, with nuclei of stressed syllables being longer than nuclei of unstressed syllables. However, the earliest studies into this matter point to the difficulty of establishing such a relation, a difficulty stemming from the fact that most unstressed syllables in English contain a schwa, that is a vowel resulting from phonological neutralization of contrasts in unstressed position. Thus, the durational differences between stressed and unstressed vowels in English are often not due to the phonetic, durational effects of stress, but rather due to phonology, which was often overlooked. For example, Lieberman (1960) reports stressed vowels of a word being longer than unstressed vowels in 66% of utterances. The test words of this particular study, however, included words such as *minute* [noun] with the second vowel being a schwa, as well as words such as *contract* [noun] with the second syllable containing a full vowel. That there is a durational difference in words of the first type is rather unassailable, and whether there is one in the words of the second type is impossible to tell, since the results were pooled together. An earlier classic study, Fry (1955) also has words like *digest* (no schwa) and *object* [V] with a schwa, treated alike. The effects of phonetic and phonological vowel reduction were differentiated in a study by Fourakis (1991), who found that sentence stress did have an influence on vowel duration.

On the perceptual side, there have been attempts to assess the importance of durational, as opposed to spectral, characteristics of vowels for the perception of stress. To exclude the influence of the formant structure on the perception of stress, Fry (1958) conducted a study in

⁵⁷ As for the manner of articulation, vowels increase in duration when followed by: voiceless stops, voiceless fricatives, nasals, voiced stops, voiced fricatives (House & Fairbanks 1953). The results regarding the role of place of articulation are inconclusive, with the results of House and Fairbanks (1953: 108) showing that vowels preceded by /d/ or /t/ are longer than the same vowel preceded by /b, g/ or /p, k/ respectively; and the results of Lehmann and Heffner (1943: 212) suggesting that they are shorter.

which formant structures were controlled for in synthesized vowels, and durational cues were still found to be important for the perception of stress. Similarly, Morton and Jassem (1965) also attested to duration of being one of the correlates of stress, however, their study, just as that of Lieberman (1960) and Fry (1955; 1958), suggests that the importance of the fundamental frequency is greater than that of duration.

Now as far as foot length is concerned, vowels in shorter feet have been found to be longer than vowels in longer feet. This phenomenon is seen as a manifestation of the purported stress-timing in English, with the shortening of nuclei of longer feet minimizing the overall durational difference between longer and shorter feet (Abercrombie 1964; White 2002). Syntax also seems to influence the duration of vowels. It has been reported (Rakerd, Sennet & Fowler: 171) that vowels are longer between major syntactic boundaries than within them, a phenomenon referred to as domain-final lengthening, as well as that vowels are shorter in longer sentences as compared to shorter sentences. This influence was not investigated in the study reported on here.

The effects of speech tempo are unsurprising in that higher speech tempo results in a decrease in duration of vowels (Lehiste 1970; Harris 1975, 1978; Gay 1978; Fourakis 1991). It is the influence of tempo on quality, and therefore the relationship between temporal and spectral reduction, that remains more controversial, as presented in the following section.

4.5.2.2 The relationship between the variation in vowel duration and vowel quality

The following possibilities for the relationship between durational and qualitative variation are conceivable:

- I. A variation in one feature automatically entails the variation in the other
 - a. Variation in duration entails variation in quality
 - b. Variation in quality entails variation in duration
- II. The conditioning factors affect the two features independently, but always in the same direction
- III. The conditioning factors affect the two features independently

A number of attempts have been made to account for the relationship between durational and spectral variation in vowels, mainly with the aim to explain the mechanism behind vowel reduction. Reduced vowels had been observed to be shorter in duration and reduced spectrally, and the question was whether the durational reduction causes spectral reduction, or whether the two kinds of reduction are motivated independently.

According to Linblom's (1963; 1990) hypo- and hyperarticulation model, decrease in duration causes articulatory and acoustic undershoot, which reflects the speakers knowledge of what kind of signal a listener needs to decode the message. Thus, spectral reduction results from a failure to reach the articulatory targets under time constraints, which results in the increase in the degree of coarticulation of the reduced vowel and its neighboring sounds. This is an example of an instantiation of possibility I.a. above. To my knowledge, no-one has proposed a mechanism which instantiated possibility I.b. Indeed, in the context of reduction under increased tempo, durational reduction is uncontroversially an automatic result of increasing tempo, and it would seem far-fetched to stipulate that it is mediated through spectral reduction

The possibility under II. would be indistinguishable from those in I.a. and I.b. in studies measuring the properties of the acoustic signal. To show that II., rather than I., is correct, neurolinguistic experiments would be needed.

A view opposing Lindblom (1963) is advocated by Harris (1975; 1978). She observes that the logical conclusion of Linblom's (1963) 'undershoot model' is that spectral reduction should accompany temporal reduction in exactly the same way, no matter what the cause behind the latter is. Having found that tempo, stress, and coda voicing cause variation in quality which is independent from variation in duration, she concludes that spectral reduction does not simply follow from temporal reduction, but that these two are independently controlled. One of the instances where durational reduction does not cause spectral reduction is in the case of vowels closed by consonants which are either voiced or voiceless. These vowels differ in duration, but not in quality (Harris 1978: 359). Thus, this view represents the possibility under III., that is the independent influence of conditioning factors on durational and spectral variation.

The correctness of the scenario in III. is also supported by Gay (1978) and Fourakis (1991), who both found that faster tempo results in shorter, but not necessarily spectrally reduced vowels, as well as by van Bergem (1993), who found that the covariation between durational and qualitative reduction is frequent, but not necessary, and by Labov (1994 [2010]: 173), who attests to instances of a negative correlation.

Summing up, although durational and qualitative variation often coincide, there does not seem to be an automatic relation between the two. In cases of reduction, for example, shortening and qualitative reduction might, and often do, coincide, but since these two kinds of reduction are independently controlled, this coincidence can be suspended. The following is the investigation of the presence or lack of such a coincidence in PDE.

4.5.2.3 Analysis

In the following analysis of the relation between the variation in quality and variation in duration, the same corpus of data was used which was employed in the investigation into the relative importance of spectral versus durational cues for the maintenance of contrasts between vowels maintained by length in Middle English, reported on in Section 4.5.1 above.

Qualitative and durational variation caused by tempo, style, coda voicing and foot length was investigated. For the analysis of the effects of tempo and style presented in 4.5.2.4 and 4.5.2.5 below, the values for duration and for F2 for the respective reading tasks were compared, yielding either a shortening, or a lengthening for duration, and a decrease or an increase in terms of F2. The data could thus be laid out in form of contingency tables where the relation between the variables could be assessed by means of a χ -square test. For the analysis of the effects of coda voicing and foot length presented in 4.5.2.6 and 4.5.2.7 below, the results of pairwise t-tests, in form of their p-values, are presented.

4.5.2.4 The effect of tempo

Table 9 below is a contingency table enabling the investigation of a relation between the variation in duration and variation in F2 as brought about by the change of reading speed (the ‘fast’ reading condition as opposed to ‘normal’ reading condition).

	smaller F2	greater F2	
shortened duration	41 (42,7%)	27 (28,1%)	68 (70,8%)
lengthened duration	15 (15,6%)	13 (13,5%)	28 (29,2%)
	56 (58,3%)	40 (41,7%)	N = 96

$$\chi^2 = 0.3688, p = 0.5437, w = 0.062, \alpha = 0.05, \beta = 0.2, N = 96$$

Table 9: Variation in duration and in F2 caused by reading speed

As can be seen in the table, the test did not yield a significant result, and the effect size is small (smaller than 0.1, the customary Cohen 1992 value of small effect size). Taking into account this very small effect size, one can tentatively assume that there is no relation between the way duration changes and the way in which F2 changes under the influence of reading speed.

4.5.2.5 The effect of style

Table 10 below is a contingency table enabling the investigation of a relation between the variation in duration and variation in F2 as brought about by the change of reading style (the ‘extra careful’ reading condition and the ‘normal’ reading condition).

	smaller F2	greater F2	
shortened duration	32 (16,7%)	55 (28,6%)	87 (45,3%)
lengthened duration	25 (13%)	80 (41,7%)	105 (54,7%)
	57 (29,7%)	135 (70,3%)	N = 192

$$\chi^2 = 3,8355, p = 0.05018, w = 0.141, \alpha = 0.05, \beta = 0.2, N = 192$$

Table 10: Variation in duration and in F2 caused by reading style

As can be seen in the table, the test did not yield a significant result, and the effect size is small (lying between 0.1 and 0.3, the customary values of small and middle effect sizes,

respectively) Taking into account this very small effect size, one can tentatively assume that there is no relation between the way duration changes and the way in which F2 changes under the influence of reading style.

4.5.2.6 The effect of coda voicing

For the analysis of the influence of coda voicing, the values for the four pairs, i.e. *bead* vs. *beat*, *bid* vs. *bit*, *bed* vs. *bet* and *bad* vs. *bat*, pooled from three reading tasks, namely ‘normal’, ‘fast’ and ‘extra careful’, separately for each gender. Female and male speakers could not be lumped together, because then the differences between durations were not normally distributed.

Table 11 below presents the results of pairwise t-tests for the respective pairs for female speakers. As can be seen, the vowels closed by voiceless consonants are shorter than those closed by voiced consonants, and this effect has reached statistical significance for all pairs. As far as F2 goes, its values were lower, suggesting a more central realization, for *bit*, *bet* and *bat*; but not for *beat*, where no statistical significance was reached.

	duration	p-value	F2	p-value
<i>bead</i> vs. <i>beat</i>	<i>bead</i> longer	$p = 6.043 \times 10^{-12}$	no difference	$p = 0.117$
<i>bid</i> vs. <i>bit</i>	<i>bid</i> longer	$p = 8.485 \times 10^{-9}$	higher for <i>bid</i>	$p = 0.024$
<i>bed</i> vs. <i>bet</i>	<i>bed</i> longer	$p = 1.521 \times 10^{-7}$	higher for <i>bed</i>	$p = 0.017$
<i>bad</i> vs. <i>bat</i>	<i>bad</i> longer	$p = 5.826 \times 10^{-7}$	higher for <i>bad</i>	$p = 0.028$

pairwise t-test, $\alpha = 0.05$, $N = 16$ ⁵⁸

Table 11: Effect of coda voicing (female speakers)

Table 12 presents the results of pairwise t-tests for the respective pairs for male speakers. As can be seen, the vowels closed by voiceless consonants are shorter than those closed by voiced consonants, and this effect has reached statistical significance for all pairs. As far as F2 goes, its values were lower, suggesting a more central realization, for *beat* and *bet*; but not for *bit* and *bat*, where no statistical significance was reached.

⁵⁸ To ensure the validity of the pairwise t-tests with $N < 30$, the differences between values were first tested for normal distribution by means of the Shapiro-Wilk test.

	duration	p-value	F2	p-value
<i>bead</i> vs. <i>beat</i>	<i>bead</i> longer	$p = 4.35 \times 10^{-7}$	higher for <i>bead</i>	$p = 3.61 \times 10^{-3}$
<i>bid</i> vs. <i>bit</i>	<i>bid</i> longer	$p = 8.3 \times 10^{-5}$	no difference	$p = 0.196$
<i>bed</i> vs. <i>bet</i>	<i>bed</i> longer	$p = 1.83 \times 10^{-4}$	higher for <i>bed</i>	$p = 0.02$
<i>bad</i> vs. <i>bat</i>	<i>bad</i> longer	$p = 4.91 \times 10^{-6}$	no difference	$p = 0.144$

pairwise t-test, $\alpha = 0.05$, $N = 16$

Table 12: Effect of coda voicing (male speakers)

Overall, vowels closed by voiced consonants are longer than vowels closed by unvoiced consonants, but not always is there a significant difference in F2.

The differences in the extent of the influence of coda voicing on duration on the one hand, and on F2 on the other can be further illustrated with the two following figures. For female speakers, tokens of *beat* were consistently shorter than tokens of *bead* (Figure 11). At the same time, the effect of coda sonority on F2, if present at all, must be a lot smaller, since nothing even approaching such a clear-cut division is apparent in Figure 12. This is in line with Harris (1978).

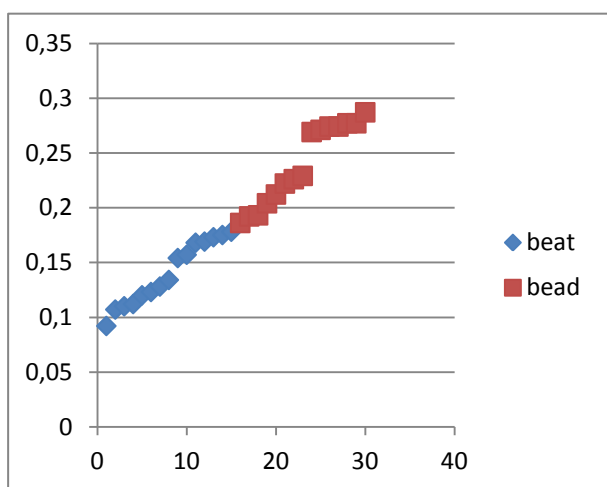


Figure 11: *beat* and *bead* tokens, ranked by vowel duration (sec)

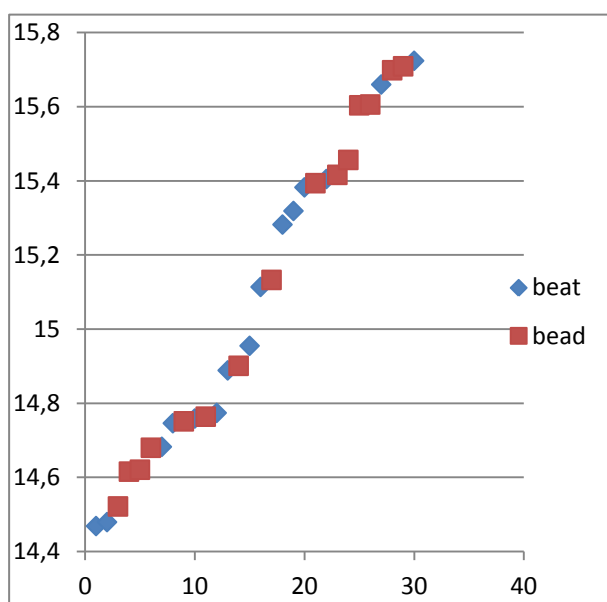


Figure 12: *beat* and *bead* tokens ranked by F2 (Bark⁵⁹)

4.5.2.7 The effect of foot length

For the analysis of the effect of foot length on durational and qualitative variation, the mean values in two-syllable words, that is *beating*, *bidding*, *betting* and *batter*, were compared against the values in one-syllable words, that is *beat*, *bid*, *bet* and *bat*, pooled from three reading tasks, namely ‘normal’, ‘fast’ and ‘extra careful’, separately for each gender. Again, female and male speakers could not be lumped together, because then the differences between durations were not normally distributed.

Table 13 below presents the results of pair-wise t-tests for the respective word pairs for female speakers. As can be seen, the vowels in two-syllable words are shorter than vowels in one-syllable words, and this effect has reached statistical significance for all pairs. As far as F2 goes, its values were lower, suggesting a more central realization, for *bidding* and *betting*, but not for *beating* and *batter*, where no statistical significance was reached.

⁵⁹ This figure employs F2 values converted to Bark scale using the equation given in Traunmüller (1990).

	duration	p-value	F2	p-value
<i>beat vs. beating</i>	<i>beating</i> shorter	$p = 1.45 \times 10^{-6}$	no difference	$p = 0.675$
<i>bid vs. bidding</i>	<i>bidding</i> shorter	$p = 5.34 \times 10^{-10}$	higher for <i>bid</i>	$p = 5.44 \times 10^{-3}$
<i>bet vs. betting</i>	<i>betting</i> shorter	$p = 2.8 \times 10^{-5}$	higher for <i>bet</i>	$p = 0.022$
<i>bat vs. batter</i>	<i>batter</i> shorter	$p = 1.3 \times 10^{-5}$	no difference	$p = 0.074$

pairwise t-test, $\alpha = 0.05$, $N = 16$

Table 13: Effect of foot length (female speakers)

Table 14 below presents the results of pair-wise t-tests for the respective word pairs for male speakers. As can be seen, the vowels in two-syllable words are shorter than vowels in one-syllable words, and this effect has reached statistical significance for all pairs. As far as F2 goes, however, the differences did not reach statistical significance for any of the pairs.

	duration	p-value	F2	p-value
<i>beat vs. beating</i>	<i>beating</i> shorter	$p = 0.021$	no difference	$p = 0.625$
<i>bid vs. bidding</i>	<i>bidding</i> shorter	$p = 7.62 \times 10^{-9}$	no difference	$p = 0.07$
<i>bet vs. betting</i>	<i>betting</i> shorter	$p = 0.017$	no difference	$p = 0.157$
<i>bat vs. batter</i>	<i>batter</i> shorter	$p = 0.018$	no difference	$p = 0.333$

pairwise t-test, $\alpha = 0.05$, $N = 16$

Table 14: Effect of foot length (male speakers)

4.5.2.8 Summary

To recapitulate, tempo and style caused variation in F2 and duration in ways which have not been found to depend on one another. As far as coda voicing and foot length are concerned,

they both have a systematic effect on vowel duration, but they influence F2 in a more complex fashion.

4.5.2.9 Discussion

From OE through ME, the stipulated correlatation where longer duration meant more peripheral, and shorter duration meant less peripheral was arguably simple enough to form a simple generalization, since there was one principal phonological factor behind the duration of a vowel, namely its phonological length (the influence of the factors mentioned here is summarized in Table 15 below). The effects of stress on duration from OE to EModE were weaker than later, since the higher predictability of stress did not call for a strong stress. It is difficult to speculate on the effects of the voicing of the coda consonant, a phenomenon attested only as late as the twentieth century, but since effects are largely visible in monosyllables only, even if one wanted to backdate it, one could not go further back than to the times of schwa loss. Thus, individual phonemes did not vary much in duration, which was largely determined by them being either phonologically long or short. This simple relation does not hold in PDE. Nowadays, longer duration sometimes does mean more peripheral, but it is not always the case. As a result, the learner perceives differences in quality as random (from the perspective of the system). Prosodic and phonetic factors cause variation in duration which is systematic. They also cause variation in quality, but a more complex one (sometimes agreeing with that in duration and sometimes not). This variation cannot be systematically accounted for by listeners, and so it is interpreted as arbitrary and intended, and so invites change.

	OE	ME	PDE
Phonological properties of the vowel itself	Yes (decisive, length)	Yes	Yes (but: secondary to tenseness)
Stress	As long as stress highly predictable, then rather small influence	Increasingly visible	Yes
Coda sonority	No	No (earliest time when it could be important: after schwa loss)	Yes
Tempo	?	?	Yes
Style	?	?	Yes
Foot length	?	?	Yes

Table 15: Some factors influencing vowel duration

In PDE different factors have a different influence on duration and on quality (F2), so that for an individual phoneme, the relationship is so complex that listeners cannot factor out the qualitative variation and therefore closely attend to (and store) the precise phonetic qualities of the vowels they are exposed to. In the historical split, where length was replaced by tenseness, this meant that the long and short counterparts of the same vowel quality went their own ways. They were separate phonemes to begin with, now they became even more distinct phonetically than they used to be. The situation nowadays is that for each individual vowel quality (and they do not have phonemic counterparts with the same quality but different length now), its duration is highly variable, contingent on a number of factors, but, crucially, its ‘inherent’ duration as an expression of phonemic length is not one of them. Its quality is also very variable. These two kinds of variation, quantitative and qualitative, do not always go hand in hand.

4.5.3 A brief cross-varietal comparison

It is worthwhile to compare the developments presented so far, which deals with the ancestral history of South-Eastern British English with other varieties of English. Below, interesting parallels to New Zealand English and to Australian English are briefly commented on. As is usual in cross-varietal comparisons, ‘standard lexical sets’ of Wells (1982) are used to

identify the relevant vowels which can be seen as counterparts of each other in the varieties under comparison due to residing in approximately the same group of lexical items.

The likelihood of the unreliability of length distinctions as a trigger for the GVS is supported by the events currently unfolding in New Zealand English. The chain shift involving front short vowels have now started to affect the ‘long’ high front FLEECE vowel. It is actually the raising DRESS vowel which is encroaching on FLEECE, and, despite the presumed length difference, DRESS is ousting FLEECE, which is in turn becoming diphthongal. Gordon et al. (2008: 41) put it down to the unreliability of length as a result of the influence of coda sonority on duration .

A possible counter-example is the case of Australian English, where new pairs established by means of length, or duration, seem to have emerged as a result of the re-alignment of vocalic qualities (Cox & Palethorpe 2007). Specifically, the clearest case is that the START and STRUT vowels seem to be qualitatively identical now, the contrast apparently being that between /ɜ:/ and /ɐ/. Additionally, the frequent monophthongization of the SQUARE and NEAR diphthongs, leaves them as vowels qualitatively overlapping with DRESS and KIT respectively, thus yielding /ɜ:/ vs. /e/ and /ɪ:/ vs. /ɪ/. The third case considered by Cox and Palethorpe (2007), that of FLEECE and KIT is least persuasive, as the former is also marked by an inglide, so the contrast is not maintained by duration alone. Also, it is incoherent to claim a duration-based contrast for both NEAR / KIT and FLEECE / KIT. A three-way contrast cannot be maintained by means of one feature. The diphthongization of FLEECE could even potentially be seen as a chain shift reaction to the raising of KIT, not unlike the diphthongization of FLEECE in New Zealand English brought about by the raising of DRESS, as argued for by Gordon et al. (2008) above.

4.6 SUMMARY

To recapitulate, these are the essential stages of the scenario proposed above.

As a result of a number of developments which affected stress assignment and the increased degree of stress-timing, phonetic duration has ceased to be a reliable expression of

phonological length contrasts. This was the case, because it has become contingent on other factors, such stress, foot length, and, recently, coda sonority

As the classes of long and short vowels were no longer reliably kept apart by means of duration, the small qualitative differences between short and long vowels have been exapted as the expression of the relevant lexical contrasts.

This exaptation of qualitative contrasts which (a) increased the number of distinct vowel qualities (b) by increasing the distance between the previously short and long counterparts initiated a movement in the vowel space and (c) resulted in an increased attention paid by listeners/speakers to minute qualitative differences, has set the vowels in motion. As there were simply many more distinct phonemic vowel qualities, and as phonetic vowel duration has become rather constrained, listeners paid more attention to qualitative differences, exaggerated them for social purposes, which set of all sorts of movements.

The Great Vowel Shift took place as a result of the exaptation of qualitative differences for the expression of lexical contrasts. The increased role of minor qualitative differences as such, together with the greater number of qualitatively distinct vowel phonemes, have persisted through the GVS, and they are a fertile breeding ground for the continued appearance of vowel shifts.

The ramifications of this account and its plausibility in light of some typological observations are presented in the following chapter.

5 CONCLUSION

5.1 SUMMARY

This thesis has attempted to address questions raised by the prevalence of vowel shifting in PDE. For that purpose, the emergence of shifting was placed within the larger context of the development of the English phonological system, and both the well-studied GVS itself as well as related vocalic changes were accounted for from an evolutionary perspective, while the formalism developed in OT was drawn upon in order to provide the descriptive framework. Empirically speaking, the account was informed by two studies: a meta-survey based on extant accounts, tracing the rate of vocalic shifts in English, and an empirical acoustic production study probing both the relative importance of quality vs. duration in maintaining vocalic contrasts in English, as well as the interrelationship between qualitative and quantitative variation in PDE. The results of the meta-study have turned out to be consistent with the assumption that the number of vowel quality changes, the building blocks of vowel chain shifts, has indeed significantly increased around the beginning of the Modern English period. The results of the empirical study confirm that quality, rather than quantity, has become the primary cue for distinguishing vowel classes in English, and they also suggest that qualitative variation is not easily predictable from quantitative variation. The latter finding gives reason to believe that the small, but nevertheless noticeable qualitative differences in phonetic vowel quality, for which listeners could not easily identify systemically rooted conditional factors, may have been interpreted as being intended and socially significant. Taking an evolutionary perspective turned out to be particularly fruitful in this investigation, as it allowed the incorporation of structural, functional and social factors into a holistic account of the long term reorganization of the English phonological system. By approaching language as an evolutionary system, such an account could be produced without taking recourse, explicitly or implicitly, to teleological explanations, and without having to overestimate the role of speakers as active agents in language change.

The account presented in this thesis sketched the following scenario: fixed lexical stress, a feature inherited from Germanic, was an important influence on long term developments in the inventory of English segments: on the one hand, it facilitated the expression of trochaic

rhythm by enabling the entrenchment of differences between stressed and unstressed syllables in terms of their segmental structures. On the other hand, it increased the frequency of stress clashes and lapses, which came to be ‘repaired’ by rhythmic adjustments on the post-lexical level. A consequence of this was that the duration of vowels became highly variable, no longer being primarily controlled by phonological vowel length, so that vowel quality differences took over as the primary cue distinguishing two classes of vowels in English, as the expression of the phonological feature $[\pm\text{tense}]$. This, in turn, had the effect of increasing the number of qualitatively distinct vowels. Since their phonetic expressions remained nevertheless variable in ways that could not systematically be attributed to factors in their phonological context, vowel shifts set in. From that point on, a high rate of vowel chain shifts has continued to be observable: as lexical contrasts among vowels have continued to be mostly preserved, individual quality changes have repeatedly been followed by others, amounting to patterns of chain shifting.

This increased rate of chain shifting appears to have remained characteristic of most PDE varieties. There are still numerous distinctive vowel qualities, and qualitative variation, not being easily predictable from quantitative variation, is salient, so that very small qualitative differences are paid attention to. Furthermore, stress, particularly in inflectional morphology, is still relatively immobile (though not as immobile as in OE, since PDE contains derivational affixes which do shift stress, i.e. that eliminate the primary stress of one form altogether). Therefore, stress and rhythm continue to induce phonetic lengthenings and shortenings of vowels. To the extent that these are rhythmically (or more generally speaking: contextually) conditioned, however, durational differences among vowels are not readily available to be given social significance. In short, since duration remained entangled in the expression of factors other than the intended lexical identity of vowels, and since the social need to express identity represents a human constant which requires phonetically variable material to draw on, it is properties relating to quality, rather than duration that were recruited for that purpose. Therefore, qualitative vowel shifts continue to be prevalent in English.

In order to test the plausibility of the proposed scenario, three of the assumptions on which it rests were checked against data. First, the assumption that the rate of vowel shifting has increased English was investigated by means of a meta-study of vocalic changes reported

for English in the extant literature. A statistical analysis attests to a significant increase in the number of the changes which can be seen as forming part of vowel chain shifts, namely unconditioned qualitative changes. Second, the assumption that vowel quality, rather than duration, is the primary cue for maintaining vowel contrasts in English, which is widely reported in the literature, was given further validation by the empirical study conducted for this thesis. Third, the data from the empirical study was further analyzed with regard to the predictability of qualitative variation from quantitative variation. The results of the analysis are consistent with the assumption that qualitative and quantitative variation are related in such complex a manner, that listeners might fail to recognize it as a systematic co-variation at all, which might result in their attributing all the more significance to minute qualitative differences.

5.2 TAKING STOCK OF A FEW GENERAL INSIGHTS

Compared to previous research on the GVS, what may be most original contribution of this thesis is that it cast the question about its inception in a new light. Instead of asking why the GVS started, the question asked here was whether, and if so why, English has become a vowel shifting language. Thus, instead of attempting to identify the reasons for the onset of a particular chain shift, the perspective was shifted, and the question asked was whether it can be argued that English has become a language in which vowel chain shifts are to be expected. This question has turned out to be meaningful and addressable in a fruitful manner.

Furthermore, I hope to have demonstrated that some of the issues which this thesis has account for could be addressed in rather satisfactory ways by the adoption of the evolutionary perspective, which appears to have provided a good vantage point from which to conceptualize sound change, and language change, in general.

Thus, one of the issues that is notoriously difficult to tackle is that of causality in language change. Within the context of the GVS, this question emerges in relation to the possible ways in which the change in quality of a given vowel can be said to cause the change in quality of another vowel. Three possible ways of dealing with it can be discerned. First, it can be said that one change causes another because speakers seek to avoid merger (Martinet 1952). Second, it can be said that individual changes, though related to each other

by forming a pattern, are not related causally (Lass 1988). Third, their relatedness can be denied altogether (Stockwell & Minkova 1988b). Since EModE vowels were often becoming increasingly similar to each other, yet did not merge but were kept distinct through shifts,⁶⁰ it seems very tempting indeed to assume some causal relation between the fact that they did not merge and the fact that shifts occurred. Yet, as Lass (1980 [2009]) points out, accounts which regard speaker agency as centrally involved in establishing the causality typically have to endow speakers with some foresight that they are unlikely to have, or with the ability to undo mergers once they have happened.

The matter of the role of functional considerations appears in a new light when seen from the evolutionary perspective. Here, it is assumed that even very small – and perceptually suboptimal – contrasts between segments can come about and exist – even for longer periods, but that segment variants whose expressions are slightly easier to keep apart perceptually may over time be transmitted more successfully than their less easily distinguishable competitors. In other words, being expressed by sounds that contrast better with sounds that express a different phoneme can act as a small bias in favor of the replication of a specific segment, and a small bias on replication can have far-reaching consequences in a longer time-scale. This idea has been applied to linguistic evolution with regard to biases relating to learnability (Smith 2011). By the same rationale of course, confusability among the expressions of two phonemes can be seen as constituting a bias against either or both of the involved constituents which are expressed by the confusable forms.

Reference to small biases on replication can prove helpful when looking at cases of system-internal causality, one which does not invoke speakers. An example which was discussed in the thesis is that of the link between prosody and segmental developments. Lass (1992a: 76) rejects it out of hand, like he rejects all functional explanations, since, according to him, neither is sufficient to explain the phenomena it purports to explain. Hence, the (ostensible, from his point of view) cases where these mechanisms are at work would have to be seen as pure and utter coincidence. However, just as the preponderance of cases of language change where functional considerations seem to be at play, the typological and diachronic data on the link between stress and its segmental effects suggest that the

⁶⁰ i.e. by laxing short vowels mergers between the classes of long and short vowels were avoided, and by the GVS and the Short Vowel Shift mergers within those classes were avoided

correlation is indeed one of causality, even though the presence of fixed lexical stress is neither a necessary nor a sufficient condition. It is enough if it is shown that the presence of a certain characteristic introduces a bias for the replication of other characteristics. The bias, however weak, if real, might give a change its directionality.

5.3 SOME OPEN QUESTIONS

There is a long tradition of accounts of the influence of prosody on segments. The attempt to put the persistence of certain kinds of changes down to the ‘articulatory setting’ or a ‘voice quality’ seems intuitively enticing. It is believable that such a property is acquired very early in the language acquisition process, and so it is conceivable that other properties of the language are under pressure to conform to it. But regardless of whether such a macro-property of a language (or language family) as the likelihood to undergo vowel shifts consists in physiology (a particular ‘articulatory setting’ depending on the presence or absence of the pitch accent) or in grammar (free vs. fixed lexical stress), the most interesting question seems to be, how such a macro-property comes into being. And actually, even the physiologically-oriented explanation boils down to an explanation pertaining to grammar in the end, as the physiological correlates of a given type of accent are not divorced from the grammatical coding of either pitch accent or another kind of accent. There must be some truth to the purported link between prosody and its segmental effects; typology and synchrony seem to converge on that point. The issue, if one is interested in the intricate details of a particular development of a particular language, as always, will be: what are the necessary and what are the sufficient conditions. Even accepting the reasonable assumption that the kind of stress a language has influences its diachronic development, if one strived to arrive at a full account of a particular instance of linguistic change, one would have to know: when does it do so, and what other properties of the language can override it. The detailed knowledge of this kind, though absolutely necessary if one wants to arrive at a full understanding of a particular development, and to make predictions about similar developments in the future, is, however, separate from the question of the kind: does stress influence the trajectory of segmental developments. The existence of counterexamples does not belie the causality of stress in historical developments, just like the existence of counterexamples to any trend does not belie the existence of that trend. What is needed is a detailed investigation of the ways in

which prosody interacts with other properties of the linguistics systems in question. To return to the issue of causality, a positive correlation between certain prosodic characteristics and certain segmental effects can point to a causal relationship. The step from (descriptively observable) correlation to (postulated) causation can be made with regard to the concept of biases on the replication process.

Naturally, allowing causality of this sort does not go a long way toward explaining how such a ‘small bias’ arises in the first place. Specifically, if it is the case that prosody can be such a prime mover behind persistent change in part due to it being acquired very early in the acquisition process, it is puzzling why such major changes in stress⁶¹ happen in the first place. This second question seems even more interesting than the first one, since even letting the nitty-gritty of the interplay of ‘forces’ of different strengths aside, it remains really puzzling how such an, if not overwhelming then certainly really powerful, force can come into existence. Is it essentially an accident, or is it itself a result of an interplay of different tendencies? For now, this has to remain an open question.

Another recalcitrant issue which is believed to benefit from the evolutionary perspective is the problem of coherence, of the GVS in particular, and of long-term changes in general. Arguably, the coherence problem can be well addressed through population thinking. When change is conceptualized as occurring to populations of linguistic constituents, then causal relations can be established for time spans larger than single generations of speakers. Specifically, to account for a vowel ‘displacing’ another vowel, reference does not have to be made to speakers who seek to ‘avoid confusion’. Instead, the following line of reasoning can be employed. When one vowel moves close to another, i.e. when many instantiations of a given phoneme give rise to expressions which lie close in the acoustic space, to the expressions of another vowel phoneme, then these two phonemes might be occasionally confused. Speakers can and do live with this. Still, over time, those variants replicate better which do the job better, i.e. those that are more different from each other. In other words, functional pressures might take time to unfold (e.g. to fight against convention which slows down change dictated by functional considerations).

⁶¹ like the change in accent in Germanic proposed by Iverson and Salmons (2003)

To conclude, it has to be admitted that answering the question if English has acquired the characteristic of ‘vowel-shiftiness’ is a daunting task, and if the answer to this question is in the affirmative, then the questions that ensue, that is why did English become a vowel shifting language, and why did it happen at that particular time in its history, are even more difficult. The contribution of the present thesis to answering these questions lies not least in trying to show that it even makes sense to address a question posed in these terms. The evidence that has been adduced here suggests that this proposal is a plausible hypothesis with regard to the issue, but at the same time it has to be acknowledged that the research program that the thesis has opened is far from completed. In particular, the following questions require further elaboration.

To begin with, the account presented here involves claims whose validity could, and therefore should, be verified by a systematic typological comparison to a representative group of languages. The first of these claims is that there is a relationship between the degree to which the rhythm of a language is stress-based, as well as the degree to which the stress in a given language is lexically fixed, and the rate of vowel shifts that this language undergoes. Likewise, the link between the degree of predictability of qualitative variation from durational variation and vowel shifting makes typological predictions. An investigation into whether the presence of a more tight-knit relation between qualitative variation and durational variation is linked to a higher rate in a given language would be a direct test of the assumption made here that it is the lack of such a tight-knit relation that contributes to vowel ‘shiftiness’.

An important residual issue concerns the status of length as a phonological feature in the phonemic system of English. Under the account provided here, length is argued to have been ousted from the vocalic system, and the structural conditions still present in English should prevent it from re-surfacing. At the same time, however, the increased influence of coda sonority on the duration of the preceding vowel seems to re-instate duration as an important phonetic cue. A long-term prediction could be that, should final stops disappear from English, which is given some credence by the fact that English final stops tend to be unreleased, new length-based vocalic contrasts could arise. Crucially, however, the role that duration might play as a cue to phonological contrasts in English now is that of enhancing the identity of consonants, not vowels, and it is primarily in the expression of vocalic identity

that duration is argued to play a diminishing role. A potential comeback of phonemic vowel length distinctions in vowels foreseen by this hypothetical scenario is not in principle ruled out by the overall account of this thesis, since it would be an accidental effect of developments in the consonantal domain, and not a result of the evolution of vowels per se. In other words, vowel lengthening before voiced codas, or, alternatively, vowel clipping before voiceless codas, clearly does not help to express the phonemic identity of a vowel, so the enhancement of this function of vowel duration does not stand in contrast to the diminishing role of duration for the expression of vowel phonemes argued for here.

The next steps towards assessing the plausibility of this account would involve testing the typological predictions that it implicitly makes, and that have been made explicit above. Furthermore, the possibility of the treatment of schwa loss in Middle English as a case of evolutionary suicide, only adumbrated here, could be put to the test, and the fruitfulness of such an endeavor could be assessed, with the application of quantitative methods developed in evolutionary biology.

To conclude, I submit this thesis with the hope that the topics that it raises and the suggestions that it makes about how they might be addressed might prove, if not true, than at least fruitful. If that is the case, I will be satisfied with the effect that my work has had.

6 LIST OF REFERENCES

- Abercrombie, David. 1964. "Syllable quantity and enclitics in English". In Abercrombie, David; Fry, D. B.; MacCarthy, P.A.D.; Scott, N.C.; Trim, J.L.M. (eds.). *In Honour of Daniel Jones*. London: Longmans, Green & Co., 216-222.
- Abercrombie, David. 1965. *Studies in phonetics and linguistics*. London: Oxford University Press.
- Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Aitchison, Jane. 2003. "Psycholinguistic perspectives on language change". In Joseph, Brian D.; Janda, Richard D. (eds.). *The handbook of historical linguistics*. Carlton: Blackwell, 736-743.
- Aitchison, Jean. 2001. *Language change. Progress or decay?* Cambridge: Cambridge University Press.
- Anderson, John R. 1993. *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Anderson, Stephen R. 1969. "An outline of the phonology of Modern Icelandic vowels". *Foundations of Language* 5(1), 53-72.
- Anderson, Stephen R. 1985. *Phonology in the twentieth century. Theories of rules and theories of representations*. (3rd edition). Chicago: University of Chicago Press.
- Anderson, Stephen R. 2000. "Reflections on "On the phonetic rules of Russian"". In Goldsmith, John A.; Laks, Bernard (eds.). *The history of phonology in the twentieth century*. *Folia Linguistica Historica* 34 (1-2), 11-27.
- Anttila, R. 1989. *Historical and comparative linguistics*. Amsterdam: John Benjamins.
- Apoussidou, Diana. 2006. *The learnability of Metrical Phonology*. PhD Thesis. University of Amsterdam.
- Archangeli, Diana. 1997. "Optimality Theory: An introduction to linguistics in the 1990s". In Archangeli, Diana; Langendoen, D. Terence (eds.). *Optimality Theory: An overview*. Oxford: Blackwell.
- Árnason, Kristján. 2011. *The phonology of Icelandic and Faroese*. New York: Oxford University Press.
- Auer, Peter. 1993. "Is a rhythm-based typology possible?". *A study of the role of prosody in phonological typology. KonTRI Working Paper No. 21*.
- Aunger, Robert (ed.). 2000. *Darwinizing culture*. Oxford: Oxford University Press.
- Awedyk, Wiesław; Hamans, Camiel. 1989. "Vowel shifts in English and Dutch: formal or genetic relation". *Folia Linguistica Historica* 8(1-2), 99-114.
- Barber, Charles. 2006. *Early Modern English*. (Reprint). Edinburgh: Edinburgh University Press.
- Baudouin de Courtenay, Jan Nieciśław. 1895. *Versuch einer Theorie phonetischer Alternationen. Ein Capitel aus der Psychophonetik*. Strassburg: Karl J. Trübner.

- Bauer, Laurie. 1979. "The second Great Vowel Shift?". *JIPA* 9(2), 57-66.
- Baumann, Andreas. 2012. "Grammaticalizing to death: evolutionary suicide of OE/ME case markers". *Paper presented at the Memetics Workshop on December, 7th, 2012 in Vienna*.
- Beal, Joan C. 2004. *English in modern times, 1700-1945*. London, New York: Oxford University Press.
- Becker, M.; Tessier, A.-M. 2011. "Trajectories of faithfulness in child-specific phonology". *Phonology* 28, 163-196.
- Beckman, Mary E. 1992. "Evidence for speech rhythms across languages". In Tohura, Y.; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.). *Speech perception, production and linguistic structure*. Tokyo: IOS Press, 457-463.
- Beddor, Patrice Speeter. 2009. "A coarticulatory path to sound change". *Language* 85(4), 785-821.
- Bermúdez-Otero, Ricardo. (in preparation). *Stratal Optimality Theory*. Oxford: Oxford University Press.
- Bermúdez-Otero, Ricardo. 2006. "Phonological change in Optimality Theory". In Brown, Keith (ed.). *Encyclopedia of language & linguistics* (2nd edition). Amsterdam: Elsevier.
- Bermúdez-Otero, Ricardo. 2007. "Diachronic phonology". In Lacy, Paul de (ed.). *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press, 497-517.
- Bermúdez-Otero, Ricardo. 2013 (forthcoming). "Amphichronic explanation and the life cycle of phonological processes". In Honeybone, Patrick; Salmons, Joseph C. (eds.). *The Oxford handbook of historical phonology*. Oxford: Oxford University Press.
- Bermúdez-Otero, Ricardo; Trousdale, Graeme. 2012. "Cycles and continua: on unidirectionality and gradualness in language change". In Nevalainen, Terttu; Traugott, Elizabeth Closs (eds.). *The Oxford handbook of the history of English*. New York: Oxford University Press, 691-720.
- Bettencourt, Luís M. A.; Lobo, José; Helbing, Dirk; Kühnert, Christian; West, Geoffrey B. 2007. "Growth, innovation, scaling, and the pace of life in cities". *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 104(17), 7301-7306.
- Blackmore, Susan J. 1999. *The meme machine*. Oxford: Oxford University Press.
- Blevins, Juliette. 2006. *Evolutionary phonology. The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Bloch, Jules (trans A. Master). 1965. *Indo-Aryan from the Vedas to modern times*. Paris: Adrien-Maisonneuve.
- Bloomfield, Leonard. 1933. *Language*. London: George Allen & Unwin.
- Bloomfield, Maurice. 1884. "On the probability of the existence of phonetic law". *The American Journal of Philology* 5(2), 178-185.

- Boeckx, Cedric; Piatelli-Palmarini, Massimo. 2005. "Language as a natural object – linguistics as a natural science". *The Linguistic Review* 22, 447-466.
- Boersma, P. 1998. *Functional phonology: formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Boersma, Paul; Weenink, David. 2009. "Praat". *Doing phonetics by computer*. Version 5.1.1.18.
- Bogacka, Anna. 2003. *Perception of second language speech sounds*. MA thesis. Adam Mickiewicz University.
- Boyland, Joyce Tang. 1996. *Morphosyntactic change in progress: a psycholinguistic approach*. PhD dissertation. University of California.
- Brighton, Henry; Smith, Kenny; Kirby, Simon. 2005. "Language as an evolutionary system". *Physics of Life Reviews* 2(3), 177-226.
- Brodie, R. 1996. *Virus of the mind. The new science of the meme*. Seattle: Integral Press.
- Brugmann, Karl. 1885. *Zum heutigen Stand der Sprachwissenschaft*. Strassburg: Karl J. Trübner.
- Bucholtz, Mary. 1999. "'Why be normal?': Language and identity practices in a community of nerd girls". *Language in Society* 28, 203-223.
- Bybee, Joan L. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, Joan; Chakraborti, P.; Jung, D.; Scheibman, J. 1998. "Prosody and segmental effect. Some paths of evolution for word stress". *Studies in Language* 22(2), 267-314.
- Campbell, Lyle. 1975. "Constraints on sound change". In Dahlstedt, K-H (ed.). *The Nordic languages and modern linguistics*, 2. Stockholm: Almqvist & Wiksell, 388-406.
- Carr, Philip; Honeybone, Patrick. 2007. "English phonology and linguistic theory: an introduction to issues, and to 'Issues in English Phonology'". *Language Sciences* 29, 117-153.
- Carstairs-McCarthy, Andrew. 1999. *The origins of complex language. An inquiry into the evolutionary beginnings of sentences, syllables, and truth*. Oxford: Oxford University Press.
- Cavalli-Sforza, Luigi L.; Feldman, Marcus W. 1973. "Cultural versus biological inheritance: phenotypic transmission from parents to children". *A theory of the effect of parental phenotypes on children's phenotypes*. *American Journal of Human Genetics* 25, 618-637.
- Ćavar, Małgorzata Ewa. 2007. "[ATR] in Polish". *Journal of Slavic Linguistics* 15(2), 207-228.
- Chambers, John K.; Trudgill, Peter. 1998. *Dialectology*. Cambridge: Cambridge University Press.
- Chen, Matthew Y. 1970. "Vowel length variation as a function of the voicing of the consonant environment". *Phonetica* 22, 129-159.

- Chen, Matthew Y.; Wang, William S-Y. 1975. "Sound change. Actuation and implementation". *Language* 51(2), 255-281.
- Chomsky, Noam. 1959. "Review of Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts". *Language* 35(1), 26-58.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass: MIT Press.
- Chomsky, Noam. 1986. *Knowledge of language. Its nature, origin, and use*. New York: Praeger Publishers.
- Chomsky, Noam; Halle, Morris. 1968 [1997]. *The sound pattern of English*. (1. MIT Press paperback. ed., 4. print). Cambridge, Mass: MIT Press.
- Clarke, Sandra; Elms, Ford; Youssef, Amani. 1995. "The third dialect of English: Some Canadian evidence". *Language Variation and Change* 7, 209-228.
- Clark, Michael J.; Hillenbrand, James M. 2003. "Quality of American English front vowels before /r/". *The Journal of The Acoustical Society of America* 33(1), 1-16.
- Clements, George N. 1985. "The geometry of phonological features". *Phonology Yearbook* 2, 225-252.
- Clements, George N. 1991. "Vowel height assimilation in Bantu languages". *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: Special Session on African Language Structures*, 25-64.
- Clements, George N. 2000. "Some antecedents of nonlinear phonology". In Goldsmith, John A.; Laks, Bernard (eds.). *The history of phonology in the twentieth century*. *Folia Linguistica Historica* 34 (1-2), 29-56.
- Clements, George N. 2006. "Feature organization". In Brown, Keith (ed.). *Encyclopedia of language & linguistics*. (2nd edition). Amsterdam: Elsevier, 433-440.
- Clements, George N.; Hume, Elizabeth. 1995. "The internal organization of speech sounds". In Goldsmith, John A. (ed.). *The handbook of phonological theory*. Cambridge: Blackwell, 245-306.
- Clements, George N.; Keyser, Samuel Jay. 1983. *CV phonology: A generative theory of the syllable*. Cambridge, Mass: MIT Press.
- Cohen, Jacob. 1992. "A power primer". *Psychological Bulletin* 112(1), 155-159.
- Cox, Felicity. 1999. "Vowel Change in Australian English". *Phonetica* 56(1-2), 1-27.
- Cox, Felicity; Palethorpe, Sallyanne. 2007. "Australian English". *JIPA* 37(03).
- Croft, William. 2000. *Explaining language change. An evolutionary approach*. Harlow: Longman.
- Croft, William. 2002. "The Darwinization of linguistics". *Selection* 3(1), 75-91.
- Croft, William. 2008. "Evolutionary Linguistics". *Annual Review of Anthropology* 37(1), 219-234.
- Cummins, Fred. 2002. "Speech rhythm and rhythmic taxonomy". In *Proceedings of Speech Prosody 2002*. Aix-en-Provence.

- Cuypere, Ludovic de. 2008. "Exploring exaptation in language change". *Folia Linguistica Historica* 26(1-2), 13-26.
- Dankovičová, Jana. 1999. "Czech". In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 70-73.
- Darwin, Charles. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray.
- Darwin, Charles. 1871. *The descent of man, and selection in relation to sex*. London: John Murray.
- Dauer, Rebecca M. 1983. "Stress-timing and syllable-timing reanalyzed". *Journal of Phonetics* 11, 51-62.
- Dauer, Rebecca M. 1987. "Phonetic and phonological components of language rhythm". In. *Proceedings of the Eleventh International Congress of Phonetic Sciences, Tallin*, 447-450.
- Dawkins, Richard. 1976 [2006]. *The selfish gene*. (2. ed., 30th anniversary ed.). Oxford: Oxford University Press.
- Dawkins, Richard. 1982. *The extended phenotype. The long reach of the gene*. Oxford: Oxford University Press.
- Dawkins, Richard. 1983. "Universal Darwinism". In Bendall, D.S (ed.). *Evolution from molecules to men*. Cambridge: Cambridge University Press, 403-425.
- Dehaene-Lambertz, G.; Houston, D. 1998. "Faster orientation latencies toward native language in two-month-old infants". *Language and Speech* 41(1), 21-43.
- Delattre, Pierre. 1962. "Some factors of vowel duration and their cross-linguistic validity". *The Journal of The Acoustical Society of America* 34(8), 1141-1143.
- Dennett, Daniel C. 1991. *Consciousness explained*. Boston: Back Bay Books.
- Dennett, Daniel C. 1995. *Darwin's dangerous idea. Evolution and the meanings of life*. New York: Simon & Schuster.
- Dercole, Fabio; Rinaldi, Sergio. 2008. *Analysis of evolutionary processes. The adaptive dynamics approach and its applications*. Princeton: Princeton University Press.
- Deumert, Ana. 2003. "Bringing speakers back in? Epistemological reflections on speaker-oriented explanations of language change". *Language Sciences* 25, 15-76.
- D'Imperio, Mariapaola; Rosenthal, Sam. 1999. "Phonetics and phonology of main stress in Italian". *Phonology* 16(1), 1-28.
- Dinkin, Aaron J. 2012. "Toward a unified theory of chain shifting". In Nevalainen, Terttu; Traugott, Elizabeth Closs (eds.). *The Oxford handbook of the history of English*. New York: Oxford University Press.
- Distin, Kate. 2005. *The selfish meme. A critical reassessment*. Cambridge: Cambridge University Press.
- Donegan, Patricia Jane. 1978. *On the Natural Phonology of vowels*. PhD thesis. The Ohio State University.

- Donegan, Patricia Jane. 1985. "The English vowel shift: a typological perspective". *Paper presented at 'The 4th International Conference on English Historical Linguistics'*. Amsterdam.
- Donegan, Patricia Jane. 1993. "Rhythm and vocalic drift in Munda and Mon-Khmer". *Linguistics of the Tibeto-Burman Area* 16(1), 1-43.
- Donegan, Patricia Jane; Stampe, David. 1979. "The study of Natural Phonology". In Dinnsen, D. A. (ed.). *Current approaches to phonological theory*. Bloomington: IUP, 126-173.
- Donegan, Patricia Jane; Stampe, David. 1983. "Rhythm and the holistic organization of language structure". In Richardson, John F.; Marks, Mitchell; Chukerman, Amy (eds.). *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Chicago: Chicago Linguistic Society, 337-353.
- Donegan, Patricia Jane; Stampe, David. 2004. "Rhythm and the synthetic drift of Munda". *Yearbook of South Asian Languages and Linguistics*, 3-36.
- Drachman, Gaberell. 1978. "Child language and language change". In Fisiak, Jacek (ed.). *Recent developments in historical phonology*. The Hague: Mouton, 123-144.
- Dresher, Elan B. 2011. "The phoneme". In van Oostendorp, Marc; Ewen, Colin J.; Hume, Elizabeth; Rice, Keren (eds.). *The Blackwell companion to phonology*. Malden, Mass: Wiley-Blackwell, 241-266.
- Dziubalska-Kończak, Katarzyna. 2002. *Beats-and-binding phonology*. Frankfurt am Main: Lang.
- Eckert, Penelope. 1988. "Adolescent social structure and the spread of linguistic change". *Language in Society* 17(2), 183-207.
- Eckert, Penelope; Mendoza-Denton, Norma. 2006. "Getting real in the Golden State (California)". In Ward, Ben; Wolfram, Walt (eds.). *American voices*. Malden, MA, Oxford: Blackwell, 139-143.
- Encrevé, Pierre. 2000. "The old and the new: some remarks on phonology and its history". In Goldsmith, John A.; Laks, Bernard (eds.). *The history of phonology in the twentieth century*. Folia Linguistica Historica 34 (1-2), 57-84.
- Engstrand, Olle. 1999. "Swedish". In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 140-142.
- Erman, Britt; Warren, Beatrice. 1999. "The idiom principle and the open choice principle". *Text* 20, 29-62.
- Fear, B. D; Cutler A.; Butterfield S. 1995. "The strong/weak syllable distinction in English". *Journal of the Acoustical Society of America* 97(3), 1893-1904.
- Fisher, Ronald Aylmer. 1930. *The genetical theory of natural selection*. Oxford: Clarendon Press.
- Fitch, W. Tecumseh. 2010. *The evolution of language*. Cambridge: Cambridge University Press.

- Flemming, Edward. 2004. "Contrast and perceptual distinctiveness". In Hayes, Bruce; Kirchner, Robert; Steriade, Donca (eds.). *Phonetically based phonology*. Cambridge: Cambridge University Press, 232-276.
- Flemming, Edward; Johnson, Stephanie. 2007. "Rosa's roses: reduced vowels in American English". *JIPA* 37(01), 83.
- Fourakis, Marios. 1991. "Tempo, stress, and vowel reduction in American English". *The Journal of The Acoustical Society of America* 90(4), 1816-1827.
- Fry, D. B. 1955. "Duration and intensity as physical correlates of linguistic stress". *The Journal of The Acoustical Society of America* 27(4), 765-768.
- Fry, D. B. 1958. "Experiments in the perception of stress". *Language and Speech* 1(1-4), 126-152.
- Garrod, Simon; Fay, Nicolas; Rogers, Shane; Walker, Bradley; Swoboda, Nik. 2010. "Can iterated learning explain the emergence of graphical symbols?". *Interaction Studies* 11(1), 33-50.
- Gay, Thomas. 1978. "Effect of speaking rate on vowel formant movements". *The Journal of The Acoustical Society of America* 63(1), 223-230.
- Gell-Mann, Murray. 1992. "Complexity and Complex Adaptive Systems". In Hawkins, J. A.; Gell-Mann, Murray (eds.). *The evolution of human languages*, 3-18.
- Gell-Mann, Murray. 1994 [2011]. *The quark and the jaguar. Adventures in the simple and the complex*. London: Abacus.
- Gess, Randall. 2003. "On re-ranking and explanatory adequacy in a constraint-based theory of phonological change". In Holt, D. Eric (ed.). *Optimality Theory and language change*. Boston: Kluwer Academic Publishers, 67-90.
- Getty, Michael. 2002. *The metre of Beowulf. A constraint-based approach*. Berlin: Mouton de Gruyter.
- Giacalone Ramat, Anna. 1998. "Testing the boundaries of grammaticalization". In Giacalone Ramat, Anna; Hopper, Paul J. (eds.). *The limits of grammaticalization*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 107-128.
- Gibbon, Dafydd. 2004. "Tone and timing: two problems and two methods for prosodic typology". In *Proceedings of the International Conference on Tonal Aspects of Languages*. Beijing, 28-30 March, 2004.
- Giegerich, Heinz J. 1992. *English phonology. An introduction*. Cambridge: Cambridge University Press.
- Giegerich, Heinz J. 2005. *Lexical strata in English. Morphological causes, phonological effects*. Cambridge: Cambridge University Press.
- Gimson, Alfred C.; Cruttenden, Alan. 2008. *Gimson's pronunciation of English*. (7th edition). London: Hodder Education.
- Goldsmith, John A. 1976. *Autosegmental phonology*. PhD Thesis. MIT.

- Gordon, Elizabeth; Hay, Jennifer; McLagan, Margaret. 2008. *New Zealand English*. Edinburgh: Edinburgh University Press.
- Gould, Stephen Jay; Vrba, Elisabeth S. 1982. "Exaptation: a missing term in the science of form". *Paleobiology* 8(1), 4-15.
- Grabe, Esther; Low, Ee Ling. 2002. "Durational variability in speech and the rhythm class hypothesis". In Gussenhoven, Carlos; Rietvelt, Tanya; Warner, Natasha (eds.). *Laboratory Phonology VII*. New York: Mouton de Gruyter, 515-546.
- Gray, R. D.; Drummond, A. J.; Greenhill, S. J. 2009. "Language phylogenies reveal expansion pulses and pauses in Pacific settlement". *Science* 323(5913), 479-483.
- Green, Antony Dubach. 1998. "The promotion of the unmarked: representing sound change in Optimality Theory". *Paper presented at the 10th International Conference on English Historical Linguistics, University of Manchester*.
- Gussenhoven, Carlos. 1999. "Dutch". In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 74-77.
- Gussmann, Edmund. 2002. *Phonology. Analysis and theory*. Cambridge: Cambridge University Press.
- Gussmann, Edmund. 2009. "Alexander J. Ellis on Modern Icelandic pronunciation". *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 126(1), 47-59.
- Halle, Morris. 1962. "Phonology in Generative Grammar". *Word* 18, 54-72.
- Halle, Morris; Clements, George N. 1983. *Problem book in phonology*. Cambridge, Mass: MIT Press.
- Halle, Morris; Keyser, Samuel Jay. 1971. *English stress. Its form, its growth, and its role in verse*. New York: Harper and Row.
- Halle, Morris; Mohanan K. P. 1985. "Segmental phonology of Modern English". *Linguistic Inquiry* 16, 57-116.
- Hamann, Silke. 2006. "On the use of Evolutionary Phonology for phonological theory". *Theoretical Linguistics* 32(2).
- Hammond, Michael. 1999. *The phonology of English. A prosodic optimality-theoretic approach*. Oxford: Oxford University Press.
- Hanulíková, Adriana; Hamann, Silke. 2010. "Slovak". *Journal of the International Phonetic Association* 40(03), 373-378.
- Harbert, Wayne. 2007. *The Germanic languages*. Cambridge: Cambridge University Press.
- Harrington, Jonathan. 2007. "Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts". In Cole, Jennifer; Hualde, J. (eds.). *Laboratory Phonology IX*. Berlin: Mouton, 125-143.
- Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
- Harris, Katherine S. 1975. "Mechanisms of duration change". In Fant, Gunnar (ed.). *Speech communication: proceedings of the Speech Communication Seminar, Stockholm, April 1-3, 1974*. Stockholm: Almqvist & Wiksell, 193-198.

- Harris, Katherine S. 1978. "Vowel duration change and its underlying physiological mechanisms". *Language and Speech* 21(4), 354-361.
- Haspelmath, Martin. 1999. "Optimality and diachronic adaptation". *Zeitschrift für Sprachwissenschaft* 18(2), 180-205.
- Haugen, Einar. 1958. "The phonemics of Modern Icelandic". *Language* 34(1), 55-88.
- Hauser, Marc D.; Chomsky, Noam; Fitch, W. Tecumseh. 2002. "The faculty of language: what is it, who has it, and how did it evolve?". *Science* 298, 1569-1579.
- Hawkins, Sarah. 2003. "Roles and representations of systematic fine phonetic detail in speech understanding". *Journal of Phonetics* 31(3-4), 373-405.
- Hawkins, Sarah. 2010. "Phonetic variation as communicative system". In Fougeron, Cécile; Kühnert, Barbara; D'Imperio, Mariapaola; Valleé, Nathalie (eds.). *Laboratory phonology X*. Berlin, New York: Mouton de Gruyter, 479-510.
- Hayes, Bruce. 1996. "Phonetically driven phonology: the role of Optimality Theory and inductive grounding". <http://roa.rutgers.edu/files/158-1196/158-1196-HAYES-0-0.PDF>.
- Hayes, Bruce; Kirchner, Robert; Steriade, Donca (eds.). 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.
- Hayes, Bruce; Steriade, Donca. 2004. "Introduction: the phonetic bases of phonological markedness". In Hayes, Bruce; Kirchner, Robert; Steriade, Donca (eds.). *Phonetically based phonology*. Cambridge: Cambridge University Press, 1-33.
- Hock, Hans H. 1992. "Causation in language change". In Bright, W. (ed.). *Oxford international encyclopedia of linguistics. Vol. 1*. London & New York: Oxford University Press, 228-231.
- Hogg, Richard M. 1992a. "Phonology and morphology". In Hogg, Richard M. (ed.). *The Cambridge history of the English language. Vol I. The beginnings to 1066*. Cambridge: Cambridge University Press, 67-167.
- Hogg, Richard Milne. 1992b. *A grammar of Old English*. Oxford: Blackwell.
- Holt, D. Eric (ed.). 2003. *Optimality Theory and language change*. Boston: Kluwer Academic Publishers.
- Hooper, Joan Bybee. 1976. "Word frequency in lexical diffusion and the source of morphophonological change". In Christie, W. (ed.). *Current progress in historical linguistics*. Amsterdam: North Holland, 96-105.
- Horn, Wilhelm; Lehnert, Martin. 1954. *Laut und Leben. Englische Lautgeschichte der neueren Zeit*. Berlin: Verlag der Wissenschaften.
- House, Arthur S. 1961. "On vowel duration in English". *The Journal of The Acoustical Society of America* 33(9), 1174-1178.
- House, Arthur S.; Fairbanks, Grant. 1953. "The influence of consonant environment upon the secondary acoustical characteristics of vowels". *The Journal of The Acoustical Society of America* 25(1), 105-113.

- Hull, David L. 1980. "Individuality and selection". *Annual Review of Ecology and Systematics* 11, 311-322.
- Hull, David L. 1988. *Science as a process. An evolutionary account of the social and conceptual development of science*. Chicago: University of Chicago Press.
- Hull, David L. 2001. *Science and selection. Essays on biological evolution and the philosophy of science*. Cambridge: Cambridge University Press.
- Hurford, James R. 2012a. "Linguistics from an evolutionary point of view". In Kempson, Ruth; Fernando, Tim; Asher, Nicholas (eds.). *Handbook of the philosophy of science: Linguistics*. Burlington: Elsevier Science, 473-498.
- Hurford, James R. 2012b. *The origins of grammar*. Oxford: Oxford University Press.
- Hyman, Larry M. 1976. "Phonologization". In Juilland, Alphonse; Devine, A. M.; Stephens, Laurence D. (eds.). *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday*. Saratoga: Anma Libri, 407-418.
- Hyman, Larry M. 1982. "The representation of length in Gokana". In *Proceedings of the First West Coast Conference on Formal Linguistics*. Stanford University, 198-206.
- Hyman, Larry M. 1985. *A theory of phonological weight*. Foris: Dordrecht.
- Hyman, Larry M. 2008. "Enlarging the scope of phonologization". *Proceedings of the Symposium on Phonologization, University of Chicago, April 25-26, 2008*.
- Iverson, Gregory K.; Salmons, Joseph C. 2003. "Laryngeal enhancement in early Germanic". *Phonology* 20(1), 43-74.
- Jacewicz, Ewa; Fox, Robert Allen; Salmons, Joseph. 2006. "Prosodic prominence effects on vowels in chain shifts". *Language Variation and Change* 18, 285-316.
- Jacobs, Haïke. 1995. "Optimality Theory and sound change". In Beckman, Jill N. (ed.). *Proceedings of the North-Eastern Linguistic Society 25 (1994)*. Amherst: GLSA Publications, 219-232.
- Jakobson, Roman. 1944 [1969]. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Frankfurt am Main: Suhrkamp.
- Jakobson, Roman; Fant, C. Gunnar M.; Halle, Morris. 1951 [1963]. *Preliminaries to speech analysis. The distinctive features and their correlates*. Cambridge, Mass: MIT Press.
- Jespersen, Otto. 1909-49. *A Modern English grammar on historical principles*. 7 vols. Copenhagen: Munksgaard.
- Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.
- Kenstowicz, Michael; Kisseberth, Charles. 1977. *Topics in phonological theory*. New York: Academic Press.
- King, Robert D. 1967. "Functional load and sound change". *Language* 43(4), 831-852.
- Kiparsky, Paul. 1965. *Phonological change*. PhD dissertation. MIT.
- Kiparsky, Paul. 1982a. *Explanation in phonology*. Dordrecht: Foris.

- Kiparsky, Paul. 1982b. "From cyclic phonology to lexical phonology". In van der Hulst, Harry; Smith, Norval (eds.). *The structure of phonological representations*. Dordrecht: Foris, 131-175.
- Kiparsky, Paul. 2000. "Opacity and cyclicity". *The Linguistic Review* 17, 351-367.
- Kiparsky, Paul. 2006. "The amphichronic program vs. Evolutionary Phonology". *Theoretical Linguistics* 32, 217-236.
- Kirby, Simon. 2002. "Natural language from Artificial Life". *Artificial Life* 8, 185-215.
- Kirby, Simon. 2007. "The evolution of meaning-space structure through iterated learning". In Lyon, Caroline; Nehaniv, Chrystopher L.; Cangelosi, Angelo (eds.). *Emergence of Communication and Language*. London: Springer, 253-268.
- Kirby, Simon; Dowman, Mike; Griffiths, Thomas L. 2007. "Innateness and culture in the evolution of language". *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 104(12), 5241-5245.
- Kirchner, Robert. 1995. "Contrastiveness is an epiphenomenon of constraint ranking". *Proceedings of the Berkeley Linguistics Society* 21, 198-208.
- Kirchner, Robert. 1997. "Contrastiveness and faithfulness". *Phonology Yearbook* 14(1), 83-111.
- Kisseberth, Charles. 1970. "On the functional unity of phonological rules". *Linguistic Inquiry* 1, 291-306.
- Labov, William. 1966 [2006]. *The social stratification of English in New York City*. (2nd edition). Cambridge: Cambridge University Press.
- Labov, William. 1981. "Resolving the Neogrammarian Controversy". *Language* 57(2), 267-308.
- Labov, William. 1994 [2010]. *Principles of linguistic change. Vol 1. Internal factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change. Vol 2. Social factors*. Malden, Mass: Blackwell.
- Labov, William; Ash, Sharon; Boberg, Charles. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter.
- Labov, William; Yaeger, Malcah Steiner Richard. 1972. *The quantitative study of sound change in progress*. Philadelphia: U.S. Regional Survey.
- Ladefoged, Peter. 1967. *Three areas of experimental phonetics*. London: Oxford University Press.
- Lahiri, A.; Riad, T.; Jacobs, Haike. 1999. "Diachronic prosody". In van der Hulst, Harry (ed.). *Word prosodic systems in the languages of Europe*. Berlin: Mouton de Gruyter, 335-422.
- Langacker, Ronald. 1987. *Foundations of cognitive grammar. Vol 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.

- Lass, Roger. 1974. "Linguistic orthogenesis: Scots vowel quantity and the English length conspiracy". *York Papers in Linguistics* 4, 7-26.
- Lass, Roger. 1976. *English phonology and phonological theory*. Cambridge: Cambridge University Press.
- Lass, Roger. 1980 [2009]. *On explaining language change*. Cambridge: Cambridge University Press.
- Lass, Roger. 1988. "Vowel shifts, great and otherwise: remarks on Stockwell and Minkova". In Kastovsky, Dieter; Bauer, Gero (eds.). *Luick revisited*. Tübingen: Narr, 395-410.
- Lass, Roger. 1990. "How to do things with junk". *exaptation in language evolution. Journal of Linguistics* 26(1), 79-102.
- Lass, Roger. 1992a. "Phonology and morphology". In Blake, Norman Francis (ed.). *The Cambridge history of the English language. Vol II. 1066 - 1476*. Cambridge: Cambridge University Press, 23-155.
- Lass, Roger. 1992b. "What, if anything, was the Great Vowel Shift?". In Rissanen, Matti; Ihalainen, Ossi; Nevalainen, Terttu; Taavitsainen, Irma (eds.). *History of Englishes: New methods and interpretations in historical linguistics*. Berlin: Mouton de Gruyter, 144-155.
- Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lass, Roger. 1999. "Phonology and morphology". In Lass, Roger (ed.). *The Cambridge history of the English language. Vol III. 1476 - 1776*(3rd edition). Cambridge: Cambridge University Press, 56-186.
- Lass, Roger. 2003. "Genetic metaphor in historical linguistics". *Alternation* 10, 47-62.
- Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, Mass: The MIT Press.
- Lehiste, Ilse. 1977. "Isochrony reconsidered". *Journal of Phonetics* 5(3), 253-263.
- Lehmann, W. P.; Heffner R-M. S. 1943. "Notes on the length of vowels (VI)". *American Speech* 18(3), 208-215.
- Leith, Dick. 2002. *A social history of English*. (2nd edition). London: Routledge.
- Lerer, Seth. 2007. *Inventing English. A portable history of the language*. New York: Columbia University Press.
- Lieberman, Philip. 1960. "Some acoustic correlates of word stress in American English". *The Journal of The Acoustical Society of America* 32(4), 451-454.
- Lightfoot, David. 2003. "Grammatical approaches to syntactic change". In Joseph, Brian D.; Janda, Richard D. (eds.). *The handbook of historical linguistics*. Carlton: Blackwell, 495-508.
- Lightfoot, David. 2007. "The ecology of languages". *D.E.L.T.A.* 23, 17-41.

- Lindblom, B. 1963. "Spectrographic study of vowel reduction". *The Journal of The Acoustical Society of America* 35(11), 1773-1781.
- Lindblom, Björn. 1990. "Explaining phonetic variation: a sketch of the H&H theory". In Hardcastle, William; Marchal, Alain (eds.). *Speech production and speech modelling*. Dordrecht: Kluwer, 403-439.
- Lisker, Leigh. 1974. "On 'explaining' vowel duration variation". *Glossa* 2, 233-246.
- Loporcaro, Michele. 2011. "Phonological processes". In Maiden, Martin; Smith, John Charles; Ledgeway, Adam (eds.). *The Cambridge history of the Romance languages. Volume 1: Structures*. Cambridge: Cambridge University Press, 109-154.
- Lotspeich, Claude M. 1921. "The cause of long vowel changes in English". *The Journal of English and Germanic Philology* 20(2), 208-212.
- Lotspeich, Claude M. 1927. "A single principle for English and Primitive Germanic sound-changes". *The Journal of English and Germanic Philology* 26(4), 467-470.
- Low, Ee Ling; Grabe, E.; Nolan, F. 2000. "Quantitative characterizations of speech rhythm: syllable-timing in Singapore English". *Language and Speech* 43(4), 377-401.
- Łubowicz, Anna. 2003. *Contrast preservation in phonological mappings*. PhD. University of Massachusetts.
- Łubowicz, Anna. 2011. "73. Chain shifts". In van Oostendorp, Marc; Ewen, Colin J.; Hume, Elizabeth; Rice, Keren (eds.). *The Blackwell companion to phonology*. Malden, Mass: Wiley-Blackwell.
- Luce, Paul A.; McLennan, Conor T. 2005. "Spoken word recognition: the challenge of variation". In Pisoni, D.; Remez, R. (eds.). *The handbook of speech perception*. Oxford: Blackwell, 591-609.
- Luick, Karl. 1896. *Untersuchungen zur englischen Lautgeschichte*. Strassburg: Trübner.
- Luick, Karl. 1921-1940. *Historische Grammatik der englischen Sprache*. Leipzig: C. H. Tauchnitz.
- Luick, Karl. 1932. "Zur neuenglischen Lautgeschichte". *Archiv für das Studium der neueren Sprachen*, 89-90.
- Luraghi, Silvia. 2010. "Causes of language change". In Luraghi, Silvia; Bubeník, Vít (eds.). *Continuum companion to historical linguistics*. London: Continuum, 358-370.
- Macaulay, George Campbell (ed.). 1900 & 1969. *The English works of John Gower. Vol. 1*. London: Oxford University Press.
- MacNeilage, Peter F. 1998. "The frame/content theory of evolution of speech production". *Behavioral and Brain Sciences* 21(4), 499-511.
- Markus, Manfred. 1994. "From stress-timing to syllable-timing: changes in the prosodic system of Late Middle English and Early Modern English". In Kastovsky, Dieter (ed.). *Studies in Early Modern English phonology*. Berlin: De Gruyter Mouton, 187-203.
- Martinet, André. 1952. "Function, structure, and sound change". *Word* 8(1), 1-32.
- Masica, Colin P. 1993. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

- Mathesius, Vilém. 1983. "Functional linguistics". In Vachek, Josef (ed.). *Praguiana. Some basic and less known aspects of the Prague Linguistic School*. Amsterdam: John Benjamins, 121-142.
- Maynard Smith, John. 1992. *Did Darwin get it right? Essays on games, sex and evolution*. New York: Chapman & Hall.
- McAllister Byun, Tara; Inkelas, Sharon; Rose, Yvan. 2012. "Transient phonology, CON and child phonological processes". *Paper given at the 20th Manchester Phonology Meeting*.
- McCarthy, John J. 1982. "Nonlinear phonology: An overview". *GLOW Newsletter* 8, 63-77.
- McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.
- McCarthy, John J. 2008. *Doing Optimality Theory. Applying theory to data*. Oxford: Blackwell.
- McCarthy, John J.; Prince, Alan 1986. "Prosodic morphology". Samizdat, Amherst and Waltham.
- McCarthy, John J.; Prince, Alan. 1995a. "Faithfulness in reduplicative identity". In Beckman, Jill N.; Walsh Dickey, Laura; Urbanczyk, Suzanne (eds.). *Papers in Optimality Theory*. Amherst, Mass.: Graduate Linguistic Student Association, 249-384.
- McCarthy, John J.; Prince, Alan. 1995b. "Prosodic morphology". In Goldsmith, John A. (ed.). *The handbook of phonological theory*. Cambridge: Blackwell, 318-366.
- McCarthy, John J.; Prince, Alan. 1999. "Faithfulness and identity in prosodic morphology". In Kager, René; van der Hulst, Harry; Zonneveld, Wim (eds.). *The prosody-morphology interface*. Cambridge: Cambridge University Press, 218-309.
- McCrohon, Luke. (to appear). "The two-stage life cycle of cultural replicators". In Waciewicz, Sławomir; Żywicznyński, Przemysław (eds.). *Theoria et historia scientiarum IX*, 151-172.
- McLennan, Conor T.; Luce, Paul A. 2005. "Examining the time course of indexical specificity effects in spoken word recognition". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(2), 306-321.
- McMahon, April. 2002. *An introduction to English phonology*. Edinburgh: Edinburgh University Press.
- McMahon, April M. S. 2000. "The emergence of the optimal? Optimality Theory and sound change". *The Linguistic Review* 17, 231-240.
- McMahon, April M. S. 2003. "When history doesn't repeat itself". In Holt, D. Eric (ed.). *Optimality Theory and language change*. Boston: Kluwer Academic Publishers, 121-142.
- McMahon, April M. S.; McMahon, Robert. 2003. "Finding families: quantitative methods in language classification". *Transactions of the Philological Society* 101(1), 7-55.
- McMahon, April M. S.; McMahon, Robert. 2005. *Language classification by numbers*. Oxford: Oxford University Press.

- Mehler, Jacques; Jusczyk, P. W.; Lambertz, G.; Halsted, N.; Bertoncini, J.; Amiel-Tison, C. 1988. "A precursor of language acquisition in young infants". *Cognition* 29, 143-178.
- Miglio, Viola; Morén, Bruce. 2003. "Merger avoidance and lexical reconstruction". In Holt, D. Eric (ed.). *Optimality Theory and language change*. Boston: Kluwer Academic Publishers, 191-228.
- Milroy, James; Milroy, Lesley. 1985. "Linguistic change, social network and speaker innovation". *Journal of Linguistics* 21(2), 339-384.
- Minkova, Donka. 1991. *The history of final vowels in English. The sound of muting*. Berlin: Mouton de Gruyter.
- Minkova, Donka. 2006. "Old and Middle English Prosody". In van Kemenade, Ans; Los, Bettelou (eds.). *The handbook of the history of English*, 94-124.
- Minkova, Donka; Stockwell, Robert P. 2003. "English vowel shifts and 'optimal' diphthongs". In Holt, D. Eric (ed.). *Optimality Theory and language change*. Boston: Kluwer Academic Publishers, 169-190.
- Moosmüller, Sylvia. 2007. *Vowels in Standard Austrian German. An acoustic-phonetic and phonological analysis*. Post-doctoral dissertation.
- Morsbach, Lorenz. 1896. *Mittelenglische Grammatik*. Halle: Max Niemeyer.
- Morton, John; Jassem, Wiktor. 1965. "Acoustic correlates of stress". *Language and Speech* 8(3), 159-181.
- Munson, B. 2007. "The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation". *Language and Speech* 50(1), 125-142.
- Murray, Robert W. 2009. "Review of Phillips, Betty S. 2006. Word frequency and lexical diffusion. New York: Palgrave Macmillan". *Language* 85(2), 487-490.
- Murray, Robert W. 2010. "Language and space: The neogrammarian tradition". In Schmidt, Jürgen Erich; Auer, Peter (eds.). *Language and space*. New York, NY: De Gruyter Mouton, 70-86.
- Myers, Scott. 1987. "Vowel shortening in English". *Natural Language & Linguistic Theory* 5(4), 485-518.
- Nazzi, T.; Bertoncini, J.; Mehler, Jacques. 1998. "Language discrimination by newborns". *Towards an understanding of the role of rhythm. Journal of Experimental Psychology: Human Perception and Performance* 24(3), 756-766.
- Nettle, Daniel; Dunbar Robin I. M. 1997. "Social markers and the evolution of reciprocal exchange". *Current Anthropology* 38(1), 93-99.
- Nguyen, Noël. 2012. "Representations of speech sound patterns in the speaker's brain: Insights from perception studies". In Cohn, Abigail C.; Fougerson, Cécile; Huffman, Marie K. (eds.). *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press, 359-368.
- Nooteboom, Sieb. 1997. "The prosody of speech: melody and rhythm". In Hardcastle, William J.; Laver, John (eds.). *The Handbook of Phonetic Sciences*, 640-673.

- Nowak, Martin A. 2006. *Evolutionary dynamics. Exploring the equations of life*. Cambridge, Mass: Belknap Press of Harvard University Press.
- Odden, David. 1986. "On the role of the Obligatory Contour Principle in phonological theory". *Language* 62(2), 353-383.
- Odden, David. 2011. "The representation of vowel length". In van Oostendorp, Marc; Ewen, Colin J.; Hume, Elizabeth; Rice, Keren (eds.). *The Blackwell companion to phonology*. Malden, Mass: Wiley-Blackwell, 465-490.
- Ohala, John J. 1974. "Experimental historical phonology". In Anderson, J. M.; Jones, C. (eds.). *Historical linguistics II. Theory and description in phonology*. Amsterdam: North Holland, 353-389.
- Ohala, John J. 1981. "The listener as a source of sound change". In Masek, Carrie S.; Hendrick, Robert A.; Miller, Mary Frances (eds.). *Papers from the parasession on language and behavior*. Chicago: Chicago Linguistic Society, 178-203.
- Pagel, Mark; Meade, Andrew. 2004. "A phylogenetic mixture model for detecting patternheterogeneity in gene sequence or character-state data". *Systematic Biology* 53, 571-581.
- Pagel, Mark; Meade, Andrew. 2005. "Mixture models in phylogenetic inference". In Gascuel, O. (ed.). *Mathematics of evolution and phylogeny*. Oxford: Oxford University Press, 121-142.
- Parvinen, Kalle. 2005. "Evolutionary suicide". *Acta Biotheoretica* 53, 241-264.
- Passy, Paul. 1890. *Etude sur les changements phonétiques et leurs caractères généraux*. Paris: Firmin-Didot.
- Paul, Hermann. 1880 [1920]. *Prinzipien der Sprachgeschichte*. (5th edition). Halle: Verlag von Max Niemeyer.
- Perkins, John. 1977. "A sociolinguistic glance at the Great Vowel Shift of English". *Papers in Psycholinguistics and Sociolinguistics. Working Papers in Linguistics* 22.
- Peterson, Godon E.; Lehiste, Ilse. 1960. "Duration of syllable nuclei in English". *The Journal of The Acoustical Society of America* 32(6), 693-703.
- Phillips, Betty S. 2006. *Word frequency and lexical diffusion*. Basingstoke: Palgrave Macmillan.
- Pierrehumbert, Janet B. 1994. "Knowledge of variation". In Beals, K.; Denton, J.; Knippen, R.; Melnar, L.; Suzuki, H.; Zeinfeld, E. (eds.). *CLS 30 Vol.2: Papers from the parasession on variation*. Chicago: CLS, 367-381.
- Pierrehumbert, Janet B. 2001. "Exemplar dynamics: Word frequency, lenition and contrast". In Bybee, Joan; Hopper, Paul (eds.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 137-157.
- Pierrehumbert, Janet B. 2002a. "Review of Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press". *Phonology* 19(3), 459-463.

- Pierrehumbert, Janet B. 2002b. "Word-specific phonetics". In Gussenhoven, Carlos; Rietvelt, Tanya; Warner, Natasha (eds.). *Laboratory Phonology VII*. New York: Mouton de Gruyter, 101-139.
- Pierrehumbert, Janet B. 2006. "The next toolkit". *Journal of Phonetics* 34(4), 516-530.
- Pike, Kenneth. 1946. *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Pinker, Steven; Bloom, Paul. 1990. "Natural language and natural selection". *Behavioral and Brain Sciences* 13(4), 707-784.
- Plotkin, Henry C. 1997. *Darwin machines and the nature of knowledge*. (2nd edition). Cambridge, Mass: Harvard University Press.
- Preston, Dennis R. 2000. "Systemic accomodation". In Britain, David; Cheshire, Jenny (eds.). *In honour of Peter Trudgill*. Amsterdam: John Benjamins, 39-58.
- Preston, Dennis R. 2008. "Diffusion and transmission: The Northern Cities Chain Shift". *Paper presented at Methods XIII, Leeds*.
- Prince, Alan. 1980. "A metrical theory for Estonian quantity". *Linguistic Inquiry* 11(3), 511-562.
- Prince, Alan. 1990. "Quantitative consequences of rhythmic organization". In Deaton, Karen; Noske, Manuela; Ziolkowski, Michael (eds.). *CLS 26-II: Papers from the Parasession on the Syllable in Phonetics and Phonology*. Chicago: Chicago Linguistic Society.
- Prince, Alan; Smolensky, Paul. 1993 [2002]. *Optimality Theory. Constraint interaction in generative grammar*.
- Prokosch, Eduard. 1939. *A comparative Germanic grammar*. Baltimore: Linguistic Society of America.
- Pulleyblank, Douglas. 1989. "Nonlinear phonology". *Annual Review of Anthropology* 18, 203-226.
- Rakerd, Brad; Sennet, William; Fowler, Carol A. "Domain-final lengthening and foot-level shortening in spoken English". *Phonetica* 44, 147-155.
- Ramus, F. 2000. "Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys". *Science* 288(5464), 349-351.
- Ramus, Franck; Nespor, Marina; Mehler, Jacques. 1999. "Correlates of linguistic rhythm in the speech signal". *Cognition* 73, 265-292.
- Rask, Rasmus. 1818. *Undersøgelse om det gamle Norske eller Islandske Sprogs Oprindelse* [Essay on the origin of the Ancient Norse or Icelandic tongue]. Copenhagen: Gyldendal.
- Reiss, Charles. 2003. "Language change without constraint reranking". In Holt, D. Eric (ed.). *Optimality Theory and language change*. Boston: Kluwer Academic Publishers, 143-168.
- Revithiadou, Anthi. 1998. *Headmost accent wins*. Doctoral dissertation. LOT.

- Rice, Keren. 2007. "Markedness in phonology". In Lacy, Paul de (ed.). *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press, 79-97.
- Ritt, Nikolaus. 1995. "Language change as evolution: looking for linguistic genes". *Vienna English Working PaperS* 4(1), 43-57.
- Ritt, Nikolaus. 1996. "Darwinising historical linguistics: applications of a dangerous idea". *Vienna English Working PaperS* 5(1-2), 27-47.
- Ritt, Nikolaus. 1997. "Mutation, variation and selection in phonological evolution: A sketch based on the case of Late Middle English a > au/_l{C/#}". In Fisiak, Jacek (ed.). *Studies in Middle English linguistics*. Berlin, New York: Mouton de Gruyter, 531-550.
- Ritt, Nikolaus. 2004. *Selfish sounds. A Darwinian approach to language change*. Cambridge: Cambridge University Press.
- Ritt, Nikolaus. 2012. "26. Middle English: Phonology". In Bergs, Alexander; Brinton, Laurel J. (eds.). *Historical linguistics of English*. Berlin: Mouton de Gruyter, 399-414.
- Roach, Peter. 1982. "On the distinction between 'stress-timed' and 'syllable-timed' languages". In Crystal, David (ed.). *Linguistic controversies: essays in linguistic theory and practice*. London: Edward Arnold, 73-79.
- Roberts, Gareth. 2008. "Language and the free-rider problem: an experimental paradigm". *Biological Theory* 3(2), 174-183.
- Rose, Nick. 1998. "Controversies in Meme Theory". *Journal of Memetics - Evolutionary Models of Information Transmission* 2(1).
- Rositzke, Harry A. 1939. "Vowel-length in General American Speech". *Language* 15(2), 99-109.
- Russel, Peter J.; Hertz, Paul E.; McMillan, Beverly. 2011. *Biology. The dynamic science*. (2nd edition). Belmont: Brooks/Cole.
- Sagey, Elisabeth Caroline. 1986. *The representation of features and relations in non-linear phonology*. PhD dissertation: MIT.
- Sakai, Katsuyuki; Hikosaka, Okihide; Nakamura, Kae. 2004. "Emergence of rhythm during motor learning". *Trends in Cognitive Sciences* 8(12), 547-553.
- Samuels, Bridget. 2011. "A minimalist program for phonology". In Boeckx, Cedric (ed.). *The Oxford handbook of linguistic minimalism*. Oxford: Oxford University Press, 574-594.
- Samuels, Michael L. 1972. *Linguistic evolution*. Cambridge: Cambridge University Press.
- Sapir, Edward. 1925. "Sound patterns in language". *Language* 1(2), 37-51.
- Saussure, Ferdinand de. 1916 [1995]. *Cours de linguistique générale*. Paris: Payot & Rivages.
- Schendl, Herbert; Ritt, Nikolaus. 2002. "Of vowel shifts great, small, long and short". *Language Sciences* 24(3-4), 409-421.
- Schenker, Alexander M. 1993. "Proto-Slavonic". In Comrie, Bernard; Corbett, Greville G. (eds.). *The Slavonic languages*. London: Routledge, 60-121.

- Schiering, René; Bickel, Balthasar; Hildebrandt Kristine A. 2012. "Stress-timed = word-based?". *Testing a hypothesis in prosodic typology. Sprachtypologie und Universalienforschung* 65(2), 157-168.
- Schleicher, August. 1863 [1873]. *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar: Hermann Böhlau.
- Schmid, Stephan. 2004. "Zur Vokalquantität in der Mundart der Stadt Zürich". *Linguistik Online* 20, 93-116.
- Schuchardt, Hugo. 1885. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: Robert Oppenheim.
- Scott-Phillips, Thomas C.; Kirby, Simon. 2010. "Language evolution in the laboratory". *Trends in Cognitive Sciences* 14(9), 411-417.
- Silverman, Daniel. 2001. "Review of Bybee, Joan. 2001. Phonology and language use. Cambridge: Cambridge University Press". *Studies in the Linguistic Sciences* 31(2), 213-216.
- Skinner, B. F. 1957. *Verbal behavior*. New York: Appleton-Century-Crofts.
- Smith, Jeremy J. 2007. *Sound change and the history of English*. Oxford: Oxford University Press.
- Smith, Kenny. 2011. "Learning bias, cultural evolution of language, and the biological evolution of the language faculty". *Human Biology* 83(2), 261-278.
- Smith, Kenny; Brighton, Henry; Kirby, Simon. 2003. "Complex systems in language evolution: the cultural emergence of compositional structure". *Advances in Complex Systems* 6(4), 537-558.
- Stampe, David. 1972. "On the natural history of diphthongs". In *Papers from the Eighth Regional Meeting of the Chicago Linguistics Society*.
- Stampe, David. 1973. "On Chapter Nine". In Kenstowicz, Michael; Kisseberth, Charles (eds.). *Issues in phonological theory*. The Hague: Mouton, 44-52.
- Stampe, David. 1979. *A dissertation on Natural Phonology*. University of Chicago.
- Steriade, Donca. 2007. "Contrast". In Lacy, Paul de (ed.). *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press, 139-157.
- Stewart, J. M. 1967. "Tongue root position in Akan vowel harmony". *Phonetica* 16, 185-204.
- Stockwell, Robert P. 1978. "Perseverance in the English vowel shift". In Fisiak, Jacek (ed.). *Recent developments in historical phonology*. The Hague: Mouton, 337-348.
- Stockwell, Robert P.; Minkova, Donka. 1988a. "A rejoinder to Lass". In Kastovsky, Dieter; Bauer, Gero (eds.). *Luick revisited*. Tübingen: Narr, 411-417.
- Stockwell, Robert P.; Minkova, Donka. 1988b. "The English Vowel Shift: problems of coherence and explanation". In Kastovsky, Dieter; Bauer, Gero (eds.). *Luick revisited*. Tübingen: Narr, 355-394.
- Sussex, Roland; Cubberley, Paul V. 2006. *The Slavic languages*. Cambridge: Cambridge University Press.

- Šuštaršič, Rastislav; Komar, Smiljana. 1999. "Slovene". In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 135-139.
- Szende, Tamás. 1999. "Hungarian". In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 104-107.
- Tesar, Bruce. 1996. "An iterative strategy for learning metrical stress in Optimality Theory". In Hughes, E.; Hughes, M.; Greenhill, A. (eds.). *The proceedings of the 21st annual Boston University Conference on Language Development, November 1996.*, 615-626.
- Tesar, Bruce; Smolensky, Paul. 1993. "The learnability of Optimality Theory: An algorithm and some basic complexity results". <http://roa.rutgers.edu/files/2-0000/2-0000-TESAR-0-0.PDF>.
- Tesar, Bruce; Smolensky, Paul. 2000. *Learnability in Optimality Theory*. Cambridge, Mass: MIT Press.
- Þráinsson, Höskuldur. 1994. "Icelandic". In König, Ekkehard; van der Auwera, Johan (eds.). *The Germanic languages*. London: Routledge.
- Torgersen, Eivind; Kerswill, Paul. 2004. "Internal and external motivation in phonetic change: Dialect levelling outcomes for an English vowel shift". *Journal of Sociolinguistics* 8(1), 25-53.
- Torgersen, Eivind; Szakay, Anita. 2011. "A study of rhythm in London: Is syllable-timing a feature of Multicultural London English?". *University of Pennsylvania Working Papers in Linguistics* 17(2).
- Traugott, Elizabeth Closs. 2004. "Exaptation and grammaticalization". In Akimoto, Minoji (ed.). *Linguistic studies based on corpora*. Tokyo: Hituzi Syobo, 133-156.
- Trautmüller, H. 1990. "Analytical expressions for the tonotopic sensory scale". *The Journal of The Acoustical Society of America* 88(1), 97-100.
- Trnka, Bohumil. 1982. "A phonemic aspect of the Great Vowel-Shift". In Bohumil Trnka. *Selected papers in Structural Linguistics. Contributions to English and general linguistics written in the years 1928-1978*. Edited by Vilém Fried. Berlin: De Gruyter Mouton, 258-262.
- Trubetzkoy, Nicolai S. 1939. *Grundzüge der Phonologie*. Prague.
- Trudgill, Peter. 1974. *The social differentiation of English in Norwich*. London: Cambridge University Press.
- Twaddell, W. Freeman. 1935. "On defining the phoneme". *Language monograph No. 16*. *Language* 11(1).
- van Bergem, Dick R. 1993. "Acoustic vowel reduction as a function of sentence accent, word stress, and word class". *Speech Communication* 12(1), 1-23.
- Vennemann, Theo. 1988. *Preference laws for syllable structure and the explanation of sound change*. Berlin: Mouton de Gruyter.
- Vincent, Nigel. 1995. "Exaptation and grammaticalization". In Andersen, Henning (ed.). *Historical Linguistics*. Amsterdam/Philadelphia: John Benjamins, 433-445.

- Wagner, Petra. 2001. "Rhythmic alternations in German read speech". In. *Proceedings of Prosody 2000*. Poznań, 237-245.
- Wang, William S-Y. 1969. "Competing changes as a cause of residue". *Language* 45(1), 9-25.
- Watson, J. D.; Crick, F. H. C. 1953. "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid". *Nature* 171, 737-738.
- Weinreich, Uriel; Labov, William; Herzog, M. I. 1968. "Empirical foundations for a theory of language change". In Lehmann, W. P.; Malkiel, Y. (eds.). *Directions for historical linguistics*. Austin: University of Texas Press, 95-195.
- Wells, J. C. 1982. *Accents of English. Vol 1. An introduction*. Cambridge: Cambridge University Press.
- White, L. S. 2002. *English speech timing: a domain and locus approach*. PhD dissertation. University of Edinburgh.
- Williams, George Christopher. 1966 [1996]. *Adaptation and natural selection. A critique of some current evolutionary thought*. Princeton, N.J: Princeton University Press.
- Wolfe, Patricia M. 1972. *Linguistic change and the Great Vowel Shift in English*. Berkeley: University of California Press.
- Wright, Sewall. 1980. "Genic and organismic selection". *Evolution* 34(5), 825-843.
- Zipf, George Kingsley. 1949. *Human behavior and the principle of least effort: an introduction to human ecology*: Addison-Wesley Press.
- Zubritskaya, Katya. 1997. "Mechanism of sound change in Optimality Theory". *Language Variation and Change* 9, 121-148.

7 APPENDIX 1: COMPARISON TO OTHER LANGUAGES

The claims put forward in this thesis need to be placed within a larger typological perspective. In this respect, at least two issues raise themselves. First, the initial observation was that English is different than, say, Romance languages, which underwent fewer unconditioned vocalic changes. Second, faced with the collapse of a duration-based length system in vowels, a development different than vowel shifting can occur, even within Germanic. A case in point is Icelandic, where a quantity-based contrast was also lost. These two issues are discussed in the following.

7.1.1 *Romance*

The following discussion of the relevant developments of Romance languages on the way from Latin is based on Loporcaro (2011). Remarkably, as he notes, none of the descendants of Latin has preserved the quantity distinctions present in the Latin vocalic system. The collapse of the length contrast, however, had very different results in different dialects. In Sardinia, the short and long counterparts simply merged, yielding a simple and symmetrical five vowel system out of the symmetrical ten vowel system of Latin. In Daco-Romance (including Romanian), an asymmetrical vowel system arose, where the back vowels merged just like in Sardinia, but in the front vowel region, short /i/ merged with long /e:/, and not with long /i:/. In the rest of Romance (including Italian, Spanish and French), the merger in the back region was the same as that in Daco-Romance, and the entire pattern was symmetrical in that the merger in the front vowels mirrored that in the back vowels. Consequently, the short high vowel /i/ merged with the long mid vowel /e:/, and not simply with its long counterpart, that is /i:/. These somehow unexpected developments, in the back vowels of Daco-Romance and both back and front vowels of common Romance, are explained by the laxness of short vowels and tenseness of long vowels already in Latin. And so it is stipulated that short /i/ was phonetically /ɪ/, and short /u/ was phonetically /ʊ/, which brought their qualities closer to those of long tense /e:/ and /o:/, and not of their original counterparts, i.e. /i:/ and /u:/. Loporcaro argues that the qualitative re-organization of

common Romance was not simultaneous with the loss of quantitative contrasts, but that the latter were brought about by Open Syllable Lengthening. In effect, the following correspondences (Table 16) between Latin and common Romance can be observed.

Latin	i:	i	e:	e	a(:)	o	o:	u	u:
common Romance	i	e	ɛ	a	ɔ	o	u		

Table 16: Latin and common Romance vocalic systems (after Loporcaro 2011: 115)

Notably, the vowel system eventually lost contrastive vowel length, but gained an additional vowel height. A number of vocalic changes followed in descendant languages, including a context-free vowel shift in French /u/ > /y/, /o/ > /u/, diphthongization of lower-mid vowels, sensitive to syllable structure and/or quality of the final vowel in most varieties, but context-free in Castilian

A major difference between the situation in Romance and in English is that the length contrast, while eventually abandoned altogether in Romance, did not disappear from English, but was replaced by another contrast, namely by a phonological tense/lax contrast. While the ten phoneme vowel system of Latin was reduced to seven phonemes in common Romance, the EModE vowel system, where the duration-based length contrast came to be replaced by a quality-based tense/lax contrast, was still twenty vowels strong, which is only two vowels fewer than the ME system and four vowels fewer than the classical OE system. Although a detailed analysis of any one lineage of Romance cannot be conducted here, it seems that indeed the number of context-free vocalic changes relative to other kinds of changes would remain lower than the relevant number traced in this thesis for English. Even though the collapse of the length contrast did arguably set in motion a number of vowel changes in Romance, this has not reached the scale found in English. Potential reasons for this difference are: (a) a lower number of vowel phonemes to begin with, (b) a greater number of monosyllables in English, (c) movable stress, that is the placement of stress on various syllables of the same root under the influence of a stress-moving suffix in Romance. As for (a), the number of phonemes was not found to correlate with the rate of change in English,

and so it does not seem to be a likely candidate responsible here. Still, it cannot be excluded that the number of distinct vowel phonemes plays a role, since the number is higher throughout the history of English, so it might be a pre-requisite, which, when joined by other relevant ingredients, results in vowel shifts. As for (b), the increased number of monosyllables in English was said to result from fixed lexical stress, and the accumulating prominence of stressed vowels at the cost of the erosion of prominence of unstressed vowels (as prominence of one element is always relative to some other element). Consequently, the number of monosyllables is unlikely to be an independent factor. As a result, (c) that is the properties of stress placement in the varieties of, or one lineage of, Romance would have to be investigated more closely to find parallels to, or interesting differences from, English.

7.1.2 Icelandic

Another development in another language with some relevance to the scenario sketched here for English is the loss of quantity-based distinctions in Icelandic. In Modern Icelandic, vowel length is predictable from the context, and so it does not mark segmental contrasts. Despite the minority view of Árnason (2011), who has recently suggested that vowel quantity is (re-)emerging as distinctive in Icelandic, most traditional accounts (e.g. Haugen 1958; Anderson 1969; Gussmann 2002), or, in Gussmann's (2009: 49) words "practically everybody working in the area of Icelandic phonetics and phonology" assumes non-contrastiveness of vowel length. Now, the vocalic system of Old Icelandic (Þráinsson 1994) was radically different, in that it was a nearly perfectly symmetrical one, where eight short vowels had their long contrastive counterparts. At any rate, the development with relevance to this thesis is the loss of distinctive vowel length on the way from Old Icelandic to Modern Icelandic. The customary division between Old and Modern Icelandic is the year 1540, the year of the publication of the first Icelandic translation of the New Testament. The knowledge of the phonology of Old Icelandic stems from the descriptions found in the First Grammatical Treatise, written in the twelfth century. The relevant changes are sketched below.

In Old Icelandic, there were eight vowel pairs distinguished by length. Before the twelfth century, there was one more such pair, namely /ε - ε:/ but it was lost after the merger of /ε/ and /e/, leaving /ε:/ without a short counterpart. The vocalic system additionally contained

three diphthongs. These are all illustrated with minimal pairs by the author of the First Grammatical Treatise.

In Modern Icelandic, vowel length is not phonemic, and all vowels (including diphthongs) can be long according to the following generalization (Gussmann 2002: 159). Vowels are phonetically long if they are (a) word final, (b) followed by a single consonant or (c) followed by a C₁C₂ cluster, where C₁ is any of [p^h, t^h, k^h, s], and C₂ is any of [j, v, r]. At the same time when distinctive length was lost from Icelandic, a number of other changes in the vocalic system took place. What happened along the way from Old Icelandic to Modern Icelandic (Þráinsson 1994) is a number of mergers, frontings, and diphthongizations. Notably, many of the Modern reflexes of the Old Icelandic long vowels are diphthongs. For the high vowels, /i/ and /u/, a diphthongal analysis is also possible (Þráinsson 1994: 146).

Remarkably, even though a number of mergers took place, not a single merger occurred in which two members of a length-based opposition fell together. This can be seen in Table 17.

i:	i	y:	y	e:	e	ø:	ø	a:	a	ɔ:	ɔ	o:	o	u:	u
í	ɪ	ɪ	ɪ	ie	ɛ	ai	œ	au	a	au	œ	ou	ɔ	u	ʏ

Table 17: Old Icelandic (top row) vowel contrasts maintained by length and their Modern Icelandic (bottom row) reflexes

The new high vowels /i/ and /u/ are actually diphthongal, a matter obscured by orthography.

As for stress in Modern Icelandic, it almost always falls on the first vowel in a word.

What seems to have happened in Icelandic, then, is that length lost its ability to mark phonological oppositions, since it came to be dependent on syllable structure. As vowels came to be lengthened in open syllables, duration could not at the same time signal the identity of the segments. Whereas in English, duration of vowels was increasingly harnessed for English feet to conform to the trochaic template and so ceased to be a reliable cue to

segmental contrasts, duration in Icelandic came to be harnessed to conform to its emerging stressed syllable structure template, a common Scandinavian process (Gussmann 2009: 50). The result in both was that vowel length disappeared as a contrastive feature from the vocalic system.

There are at least two important differences between Icelandic and English in terms of the further developments of the vowel system. First, like in Romance, the contrastive feature of length disappeared completely from Icelandic rather than being replaced by a new feature, as was the case in English. Second, more importantly, the recoding of length contrasts as qualitative contrasts did not result in vowel shifts in Icelandic. As already noted, members of the two categories have not collapsed in Icelandic, just as they did not do so in English.

Thus, fixed lexical stress combined with a prosodic requirement on the duration of stressed vowels resulted, both in English and in Icelandic in a number of context-free qualitative changes. The number of context-free qualitative changes is indeed remarkable in Icelandic. In that respect, it looks a lot more similar to English than is the case with Romance, which could be put down to the somewhat more movable stress in Romance than in Icelandic or English. What happened in Icelandic looks like the halted first stages in the major vocalic reorganization in EModE. In both cases, the short vowels fell, or laxed. In both cases, high vowels diphthongized. These changes were one of the first changes in EModE, later followed by other context-free qualitative changes, some which came to form a chain shift. One possible reason for this difference was that by the sixteenth century, when these changes started taking place, English already contained a number of diphthongs, whereas in Icelandic the ‘diphthongal space’ was free for the taking by the former long vowels.





8 APPENDIX 2: A TABLE OF VOCALIC CHANGES

5th	6th	7th	8th	9th	10th	11th	12th
i	i	i	i	i	i	i	i
i:	i:	i:	i:	i:	i:	i:	i:
ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ
ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ
ɛ:	ɛ:	ɛ:	ɛ:	ɛ:	ɛ:	ɛ:	ɛ:
e	e	e	e	e	e	e	e
e:	e:	e:	e:	e:	e:	e:	e:
æ	æ	æ	æ	æ	æ	æ	æ
æ:	æ:	æ:	æ:	æ:	æ:	æ:	æ:
a	a	a	a	a	a	a	a
a:	a:	a:	a:	a:	a:	a:	a:
ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ
ə	ə	ə	ə	ə	ə	ə	ə
ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:
ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ
ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:
ɒ	ɒ	ɒ	ɒ	ɒ	ɒ	ɒ	ɒ
ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ
ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:
o	o	o	o	o	o	o	o
o:	o:	o:	o:	o:	o:	o:	o:
u	u	u	u	u	u	u	u
u:	u:	u:	u:	u:	u:	u:	u:
ʉ	ʉ	ʉ	ʉ	ʉ	ʉ	ʉ	ʉ
y	y	y	y	y	y	y	y
y:	y:	y:	y:	y:	y:	y:	y:
ɪə	ɪə	ɪə	ɪə	ɪə	ɪə	ɪə	ɪə
ʊə	ʊə	ʊə	ʊə	ʊə	ʊə	ʊə	ʊə
ɔə	ɔə	ɔə	ɔə	ɔə	ɔə	ɔə	ɔə
<ea>	<ea>	<ea>	<ea>	<ea>	<ea>	<ea>	<ea>
<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>
<ie>	<ie>	<ie>	<ie>	<ie>	<ie>	<ie>	<ie>
<īe>	<īe>	<īe>	<īe>	<īe>	<īe>	<īe>	<īe>
<io>	<io>	<io>	<io>	<io>	<io>	<io>	<io>
<eo>	<eo>	<eo>	<eo>	<eo>	<eo>	eə	eə
<īo>	<īo>	<īo>	<īo>	<īo>	<īo>	<īo>	<īo>
<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>
ø	ø	ø	ø	ø	ø	ø	ø
ø:	ø:	ø:	ø:	ø:	ø:	ø:	ø:
eɪ	eɪ	eɪ	eɪ	eɪ	eɪ	eɪ	eɪ
aɪ	aɪ	aɪ	aɪ	aɪ	aɪ	aɪ	aɪ
uɪ	uɪ	uɪ	uɪ	uɪ	uɪ	uɪ	uɪ
oɪ	oɪ	oɪ	oɪ	oɪ	oɪ	oɪ	oɪ
/ɪu/	/ɪu/	/ɪu/	/ɪu/	/ɪu/	/ɪu/	/ɪu/	/ɪu/
aʊ	aʊ	aʊ	aʊ	aʊ	aʊ	aʊ	aʊ
ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ
eʊ	eʊ	eʊ	eʊ	eʊ	eʊ	eʊ	eʊ
oʊ	oʊ	oʊ	oʊ	oʊ	oʊ	oʊ	oʊ
əʊ	əʊ	əʊ	əʊ	əʊ	əʊ	əʊ	əʊ

9th	10th	11th	12th	13th	14th	15th	16th
I	I	I	I	I	I	I	I
II	II	II	II	II	II	II	II
i	i	i	i	i	i	i	i
i:	i:	i:	i:	i:	i:	i:	i:
ī	ī	ī	ī	ī	ī	ī	ī
ī	ī	ī	ī	ī	ī	ī	ī
ē	ē	ē	ē	ē	ē	ē	ē
ē:	ē:	ē:	ē:	ē:	ē:	ē:	ē:
e	e	e	e	e	e	e	e
e:	e:	e:	e:	e:	e:	e:	e:
æ	æ	æ	æ	æ	æ	æ	æ
æ:	æ:	æ:	æ:	æ:	æ:	æ:	æ:
a	a	a	a	a	a	a	a
a:	a:	a:	a:	a:	a:	a:	a:
Λ	Λ	Λ	Λ	Λ	Λ	Λ	Λ
ə	ə	ə	ə	ə	ə	ə	ə
ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:
ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ
ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:
ɒ	ɒ	ɒ	ɒ	ɒ	ɒ	ɒ	ɒ
ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ
ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:
o	o	o	o	o	o	o	o
o:	o:	o:	o:	o:	o:	o:	o:
u	u	u	u	u	u	u	u
u:	u:	u:	u:	u:	u:	u:	u:
ʉ:	ʉ:	ʉ:	ʉ:	ʉ:	ʉ:	ʉ:	ʉ:
y	y	y	y	y	y	y	y
y:	y:	y:	y:	y:	y:	y:	y:
ɪə	ɪə	ɪə	ɪə	ɪə	ɪə	ɪə	ɪə
ʊə	ʊə	ʊə	ʊə	ʊə	ʊə	ʊə	ʊə
ɔə	ɔə	ɔə	ɔə	ɔə	ɔə	ɔə	ɔə
<ea>	<ea>	<ea>	<ea>	<ea>	<ea>	<ea>	<ea>
<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>
<ie>	<ie>	<ie>	<ie>	<ie>	<ie>	<ie>	<ie>
<īe>	<īe>	<īe>	<īe>	<īe>	<īe>	<īe>	<īe>
<io>	<io>	<io>	<io>	<io>	<io>	<io>	<io>
<eo>	<eo>	eə	eə	eə	eə	eə	eə
<īo>	<īo>	<īo>	<īo>	<īo>	<īo>	<īo>	<īo>
<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>
ø	ø	ø	ø	ø	ø	ø	ø
ø:	ø:	ø:	ø:	ø:	ø:	ø:	ø:
er	er	er	er	er	er	er	er
ar	ar	ar	ar	ar	ar	ar	ar
ur	ur	ur	ur	ur	ur	ur	ur
or	or	or	or	or	or	or	or
/iu/	/iu/	/iu/	/iu/	/iu/	/iu/	/iu/	/iu/
au	au	au	au	au	au	au	au
eu	eu	eu	eu	eu	eu	eu	eu
ou	ou	ou	ou	ou	ou	ou	ou
əu	əu	əu	əu	əu	əu	əu	əu

13th	14th	15th	16th	17th	18th	19th	20th
i	i	i	i	i	i	i	i
i:	i:	i:	i:	i:	i:	i:	i:
ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ
ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ
ɛ:	ɛ:	ɛ:	ɛ:	ɛ:	ɛ:	ɛ:	ɛ:
e	e	e	e	e	e	e	e
e:	e:	e:	e:	e:	e:	e:	e:
æ	æ	æ	æ	æ	æ	æ	æ
æ:	æ:	æ:	æ:	æ:	æ:	æ:	æ:
a	a	a	a	a	a	a	a
a:	a:	a:	a:	a:	a:	a:	a:
ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ
ə	ə	ə	ə	ə	ə	ə	ə
ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:	ɜ:
ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ
ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:	ɑ:
ɒ	ɒ	ɒ	ɒ	ɒ	ɒ	ɒ	ɒ
ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:	ɔ:
o	o	o	o	o	o	o	o
o:	o:	o:	o:	o:	o:	o:	o:
u	u	u	u	u	u	u	u
u:	u:	u:	u:	u:	u:	u:	u:
ʊ	ʊ	ʊ	ʊ	ʊ	ʊ	ʊ	ʊ
ʊ:	ʊ:	ʊ:	ʊ:	ʊ:	ʊ:	ʊ:	ʊ:
ɥ	ɥ	ɥ	ɥ	ɥ	ɥ	ɥ	ɥ
y	y	y	y	y	y	y	y
y:	y:	y:	y:	y:	y:	y:	y:
ɪə	ɪə	ɪə	ɪə	ɪə	ɪə	ɪə	ɪə
ʊə	ʊə	ʊə	ʊə	ʊə	ʊə	ʊə	ʊə
ɔə	ɔə	ɔə	ɔə	ɔə	ɔə	ɔə	ɔə
<ea>	<ea>	<ea>	<ea>	<ea>	<ea>	<ea>	<ea>
<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>	<ēa>
<ie>	<ie>	<ie>	<ie>	<ie>	<ie>	<ie>	<ie>
<īe>	<īe>	<īe>	<īe>	<īe>	<īe>	<īe>	<īe>
<io>	<io>	<io>	<io>	<io>	<io>	<io>	<io>
eə	eə	eə	eə	eə	eə	eə	eə
<ɪo>	<ɪo>	<ɪo>	<ɪo>	<ɪo>	<ɪo>	<ɪo>	<ɪo>
<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>	<ēo>
ø	ø	ø	ø	ø	ø	ø	ø
ø:	ø:	ø:	ø:	ø:	ø:	ø:	ø:
eɪ	eɪ	eɪ	eɪ	eɪ	eɪ	eɪ	eɪ
aɪ	aɪ	aɪ	aɪ	aɪ	aɪ	aɪ	aɪ
uɪ	uɪ	uɪ	uɪ	uɪ	uɪ	uɪ	uɪ
oɪ	oɪ	oɪ	oɪ	oɪ	oɪ	oɪ	oɪ
/iu/	/iu/	/iu/	/iu/	/iu/	/iu/	/iu/	/iu/
aʊ	aʊ	aʊ	aʊ	aʊ	aʊ	aʊ	aʊ
ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ	ɛʊ
eʊ	eʊ	eʊ	eʊ	eʊ	eʊ	eʊ	eʊ
oʊ	oʊ	oʊ	oʊ	oʊ	oʊ	oʊ	oʊ
əʊ	əʊ	əʊ	əʊ	əʊ	əʊ	əʊ	əʊ

Rubric:

-  - no change
-  - conditioned change
-  - unconditioned qualitative change resulting in merger
-  - unconditioned qualitative change not resulting in merger

9 APPENDIX 3: MATERIALS USED FOR RECORDINGS FOR THE STUDY DESCRIBED IN SECTIONS 4.5.1 AND 4.5.2

1. Read the sentences slowly, stressing the words in bold. **Please leave a 2-3 second pause between the sentences.**

- ☐ I **think** that she talked to him **again**.
- ☐ I **think** that she spoke of the **beating**.
- ☐ I **think** that she talked about a **bed**.
- ☐ I **think** that she spoke of the **betting**.
- ☐ I **think** that she wants to hear this **beat**.
- ☐ I **think** that she talked about a **bat**.
- ☐ I **think** that she's champing at the **bit**.
- ☐ I **think** that she spoke of the **bidding**.
- ☐ I **think** that she said he wasn't **bad**.
- ☐ I **think** that she talked about a **bead**.
- ☐ I **think** that she wants to make a **bid**.
- ☐ I **think** that she wants him not to **bet**.
- ☐ I **think** that she said it wasn't **batter**.
- ☐ I **think** that he doesn't like her **dog**.

2. Read the sentences quickly. Imagine that your cell phone's battery is almost dead, and this could be the last sentence, and it is very important that you get it across. Still, **please leave a 2-3 second pause between the sentences.**

- ☐ I **think** that she talked to him **again**.
- ☐ I **think** that she spoke of the **beating**.
- ☐ I **think** that she talked about a **bed**.
- ☐ I **think** that she spoke of the **betting**.
- ☐ I **think** that she wants to hear this **beat**.
- ☐ I **think** that she talked about a **bat**.
- ☐ I **think** that she's champing at the **bit**.
- ☐ I **think** that she spoke of the **bidding**.
- ☐ I **think** that she said he wasn't **bad**.
- ☐ I **think** that she talked about a **bead**.
- ☐ I **think** that she wants to make a **bid**.
- ☐ I **think** that she wants him not to **bet**.
- ☐ I **think** that she said it wasn't **batter**.
- ☐ I **think** that he doesn't like her **dog**.

3. Read the following passage.

Mary and John visited their friend, Susan the other day. When they came back home, they weren't sure what Susan's little daughter was trying to tell them. 'I think that she talked about a bed', said Mary. John was not at all sure this was the case. 'I think that she talked about a bead', he replied. Also, John could not remember what Susan thought about her new friend, Winston. 'I think that she said he wasn't bad', Mary helped him out. 'And what did she say about that fishy auction?', John inquired further. 'I think that she wants to make a bid', gasped Mary, and left the room. Somewhere in the distance, a dog barked.

4. Read the word list slowly, like for a pronouncing dictionary.

bud

bet

bead

beating

bit

batter

bidding

betting

bet

bat

bidding

beat

bed

bad

bid

beat

bed

betting

bead

bit

bat

beating

batter

bid

bad

bod

10 APPENDIX 4: ABSTRACT

This thesis presents the Great Vowel Shift (GVS) and related vocalic changes from a strictly evolutionary perspective. The usual questions of why or when the GVS took place are replaced with an attempt to posit a change in the overall tendency of English to experience vowel shifts, with the GVS being only one particular instantiation. It is proposed that English has acquired a tendency to undergo vowel chain shifts as a result of a number of events in its history, beginning already with the fixing of lexical stress in Germanic, through to the advent of rhythmically motivated post-lexical stress demotions and promotions in late Middle English. A co-evolution of stress placement, rhythmically motivated adjustments and binary rhythm is argued to have resulted in the de-coupling of qualitative variation from durational variation in vowels. This, in turn, is posited to have resulted in the loss of length as a feature distinguishing vowel contrasts, and, through the rise in the role of minute qualitative differences in maintaining vocalic identity, in the increased rates of vowel shifting. The account is informed by two studies conducted for the purposes of the present thesis. First, a statistical metastudy of vocalic changes reported on in the literature was conducted, and its results are consistent with the assumption that the tendency of English to undergo vowel shifts has indeed increased. Second, an empirical acoustic study was conducted, and its results are consistent with the assumption that quality takes precedence over quantity in the expression of phonemic oppositions in English vowels, as well as with the assumption that qualitative variation is not easily predictable from durational variation in Present-Day English. Crucial stages in the development of English phonological system are represented by means of the formalism developed in Optimality Theory.

11 APPENDIX 5: KURZFASSUNG

Diese Dissertation stellt die Englische Vokalverschiebung und verwandte vokalische Veränderungen von einer evolutionären Perspektive dar. Die üblichen Fragen nach den Ursachen der Vokalverschiebung werden mit dem Versuch, den Wechsel der allgemeinen Tendenz des Englischen, Vokalverschiebung zu erleben, ersetzt, wobei die Englische Vokalverschiebung nur eine Instantiierung dieser Tendenz ist. Es wird vorgeschlagen, dass Englisch eine Tendenz, Vokalverschiebungen zu erleben, erworben hat, infolge einer Vielfalt von Ereignissen in seiner Entwicklung, von der Festlegung des Wortakzents bereits im Germanischen, bis zum Auftritt von rhythmisch motivierten postlexikalischen Akzentherabstufungen und –erhebungen im späten Mittelenglischen. Eine Koevolution von Wortakzent, rhythmisch motivierten Anpassungen und binärem Rhythmus habe als Folge eine Entkopplung von qualitativer Variation von quantitativer Variation in Vokalen. Dies, wiederum, habe als Folge den Verlust von Vokallänge, eines Merkmals das phonemische Vokalkontraste erhalten hatte, und, durch den Aufstieg der Rolle der minutiösen qualitativen Unterschiede für das Erhalten der vokalischen Identitäten, eine erhöhte Rate der Vokalveränderungen. Diese Darstellung beruht auf zwei Studien, die für diese Dissertation durchgeführt worden sind. Erstens, es wurde eine statistische Metastudie von vokalischen Veränderungen, die in der Literatur ausgewiesen worden sind, durchgeführt, deren Ergebnisse mit der Annahme, die Tendenz des Englischen, vokalische Veränderungen zu erleben sei aufgestiegen, konsistent sind. Zweitens, es wurde eine empirische, akustische Studie durchgeführt, deren Ergebnisse, mit der Annahme, Qualität habe Vorrang vor Quantität als Expression von phonemischen Oppositionen in englischen Vokalen, als auch mit der Annahme, qualitative Variation sei von qualitativer Variation im heutigen Englischen nicht leicht berechenbar, konsistent ist. Kritische Stadien in der Entwicklung des englischen phonologischen Systems werden mittels Formalismus, der in der Optimalitätstheorie entwickelt worden ist, dargestellt.

12 APPENDIX 6: CURRICULUM VITAE

Persönliche Daten

Name Kamil Kaźmierski
Geboren 03.06.1985, Inowrocław, Polen

Ausbildung

2009- Doktoratsstudium, Anglistik und Amerikanistik, Uni Wien

2007-2009 M.A. equivalentes Studium ('magister'), Adam Mickiewicz University,
Poznań, Polen (Englische Sprachwissenschaft)

2004-2007 B.A. equivalents Studium ('licencjat'), Adam Mickiewicz University, Poznań,
Polen (Englische Sprachwissenschaft)

Akademische stellen

2009-2013 wissenschaftlicher Mitarbeiter in Ausbildung, am Institut für
Anglistik und Amerikanistik, Uni Wien

Lehre

2009-2013 PS 1 – Linguistic Proseminar: Introductory Seminar
2008-2009 Sprachkompetenz – Englischlehrer
2006-2007 Sprachkompetenz – Englischlehrer in Ausbildung

Publikationen

2009 "On the role of perception in the acquisition of the *peach – pitch* contrast by
Polish learners of English". *VIEWS* 18(2), 24-41.

- 2009 “The *peach-pitch* problem – on the acquisition of the /i:/ - /ɪ/ contrast by Polish learners of English”. Adam Mickiewicz University (unpublished MA thesis).

Besuchte Konferenzen und gehaltene Vorträge

- 2012 43rd Poznań Linguistic Meeting (PLM2012), Poznań, Polen.
Vortrag: “Exaptation in the history of the English vowel system: What it really means to take a Darwinian view of language evolution.”
- 2012 17th International Conference on English Historical Linguistics (ICEHL-17), Zürich, die Schweiz.
Vortrag: “Has English become a vowel shifting language, and if so why?”
- 2012 Memetics Workshop. Wien, Österreich.
Vortrag: “The replacement of length by tenseness has been a memetic process.”
- 2012 Symposium: Studying Change in the History of English: New Directions in Linguistics. Wien, Österreich.
- 2011 44th Annual Meeting of the Societas Linguistica Europaea (SLE 2011), Logroño, Spanien.
Vortrag: “The ties between variation in vowel quality and duration.”
- 2010 41st Poznań Linguistic Meeting (PLM2010), Gniezno, Polen.
Vortrag: “The relation between variation in vowel quantity and vowel quality.”
- 2010 16th International Conference on English Historical Linguistics (ICEHL-16), Pécs, Ungarn.
Vortrag (mit Nikolaus Ritt): “The influence of morphology on the evolution of English phonotactics.”

- 2010 Language as an evolutionary system: a multidisciplinary approach, Edinburgh, Schottland.
- 2010 Faculty workshop: Facts and myths in language history, Wien, Österreich.
- 2009 37. Österreichische Linguistiktagung, Salzburg, Österreich.

Weitere wissenschaftliche Tätigkeiten

- 2010- Editorial team. *Folia Linguistica Historica*
- 2009- Editorial board. *Vienna English Working Papers (VIEWS)*

Mitgliedschaften in wissenschaftlichen Gremien

Societas Linguistica Europaea

EDV Kenntnisse

Microsoft Office, Praat, STx, R

Sprachen

Englisch, Polnisch, Deutsch